SIAM J. IMAGING SCIENCES
Vol. 10, No. 3, pp. 1416–1445

The Space of Essential Matrices as a Riemannian Quotient Manifold*

Roberto Tron† and Kostas Daniilidis‡

Abstract. The essential matrix, which encodes the epipolar constraint between points in two projective views, is a cornerstone of modern computer vision. Previous works have proposed different characterizations of the space of essential matrices as a Riemannian manifold. However, they either do not consider the symmetric role played by the two views or do not fully take into account the geometric peculiarities of the epipolar constraint. We address these limitations with a characterization as a quotient manifold that can be easily interpreted in terms of camera poses. While our main focus is on theoretical aspects, we include applications to optimization problems in computer vision.

Key words. epipolar geometry, Riemannian geometry, optimization

AMS subject classifications. 68Q25, 68R10, 68U05

DOI. 10.1137/16M1091332

1. Introduction.

The *essential matrix* and the *epipolar constraint*, introduced by [16], have been a major mainstay of compute vision for the last 30 years and are the basic building block in any Structure from Motion (SfM) system. Its robust estimation from image data is now standard course material (see the textbooks from [13] and [17]). In practical terms, an essential matrix encodes an *epipolar configuration* (i.e., an Euclidean motion between two camera views) as a matrix in $\mathbb{R}^{3\times3}$. The space of essential matrices is a subset of $\mathbb{R}^{3\times3}$, but the algebraic relations imposed by the epipolar constraint render its geometry and its relation with the space of epipolar configurations far from trivial.

There have been a few attempts to characterize the space of essential matrices as a Riemannian manifold. The earliest works in this aspect are from [22] and [18], with a follow-up from [7]; in these works, one of the two views is chosen as the global reference frame, and essential matrices are parametrized using the (normalized) relative poses between cameras (unit vectors for the translations and rotation matrices). This parametrization implies a preferential treatment of one of the two cameras and breaks the natural symmetry of the constraint. A different representation, based on the Singular Value Decomposition (SVD) of the essential matrix, has been used in several papers [9, 14, 24, 25]. While this representation has a natural symmetry, previous works do not provide an intuitive geometric interpretation of its parameters. In addition, they do not completely take into account the well-known twisted-pair ambiguity,

*Received by the editors August 29, 2016; accepted for publication (in revised form) March 27, 2017; published electronically August 31, 2017. A preliminary version of part of this work has appeared in the conference proceeding [28].
http://www.siam.org/journals/siims/10-3/M109133.html
Funding: The work of the authors was supported by grants NSF-IIP-0742304, NSF-OIA-1028009, ARL MAST-CTA W911NF-08-2-0004, and ARL RCTA W911NF-10-2-0016, NSF-DGE-0966142, and NSF-IIS-1317788.

†Department of Mechanical Engineering, Boston University, Boston, MA 02215 (tron@bu.edu, http://sites.bu.edu/tron).
‡GRASP Lab, University of Pennsylvania, Philadelphia, PA 19104 (kostas@seas.upenn.edu, www.cis.upenn.edu/~kostas).

i.e., the fact that four different epipolar configurations correspond to the same essential matrix (with an arbitrary choice of sign). Moreover, when considered, the algorithm used for the computation of the logarithm map (which is related to the notion of geodesics in this space) is neither efficient nor rigorously motivated.

In this work, we propose characterizations of the spaces of essential matrices and epipolar configurations as Riemannian quotient manifolds. We make the following contributions:

1. Our representation is related to the aforementioned SVD formulation, but we derive our results from a particular choice of the global reference frame, leading to a clear geometric interpretation of the parameters.
2. We clarify the relation between the chirality constraint (i.e., the constraint that all the points lie in front of both cameras), the space of essential matrices, and the space of epipolar configurations.
3. We use the theory of quotient manifold to endow the space of essential matrices and the space of epipolar configurations with a Riemannian manifold structure. This procedure leads to a natural characterization of geodesics, distance, and curvature from those defined in the space of rotations.
4. We provide expressions for the curvature of the manifolds, showing that it is nonnegative. This is an important fact for some optimization algorithms.
5. Our treatment includes procedures to efficiently and correctly compute the logarithm map and distance function between points on the manifolds.
6. We apply the theory to problems in two-view SfM, showing that the proposed representation provides an effective way to parametrize optimization problems and a meaningful notion of distance between epipolar configurations.

Some material in this paper might appear quite basic for any reader versed in computer vision. However, it is necessary to revisit it and place it in the context of our parametrization.

Paper outline. The paper is organized as follows. We first introduce our notation and review basic concepts in Riemannian geometry and group theory (section 2). We then derive a canonical decomposition of essential matrices that is given by a particular choice of the global reference frame (subsection 4.1), use it to characterize the space of essential matrices as a *quotient space* (subsection 4.2), and show its interpretation in terms of image vectors (subsection 4.3). Using the chirality constraint, we derive the *signed normalized essential space* from the space of essential matrices and show that it is a *quotient manifold* (subsection 5.3); we derive expressions and algorithms for computing geodesics, distances, and curvature of this manifold (subsections 5.4 to 5.6). We use these results to then go back to the space of essential matrices and show that it is also a manifold (section 6). Finally, we show an application of the theory to optimization problems in computer vision (section 7).

2. Definitions and notation.

In this section, we define the notation used in this paper and review several notions from Riemannian geometry and group theory. For the most part, these are well-established results, and we include here just the minimum necessary to follow the paper while referring the reader to the literature for complete and rigorous definitions [6, 21]. Nonetheless, subsection 2.7 includes results for $SO(3) \times SO(3)$ as a Lie group that, although simple, have never been explicitly presented before.

2.1. General notation. We denote as $I \in \mathbb{R}^{3\times 3}$ the identity matrix and as $P_z = \mathrm{diag}(1,1,0)$ the standard projector on the xy-plane. As customary, we use $\mathfrak{so}(3)$ to indicate the space of 3×3 skew-symmetric matrices. For standard vectors $a \in \mathbb{R}^3$, $[a]_\times : \mathbb{R}^3 \to \mathfrak{so}(3)$ denotes the matrix representation of the cross-product operator, i.e., $[a]_\times b = a \times b$ for all $a, b \in \mathbb{R}^3$. We use $[a]_\times^{\mathrm{inv}} : \mathfrak{so}(3) \to \mathbb{R}^3$ to denote the inverse of this linear mapping. We use $\mathrm{sym}(A)$ and $\mathrm{skew}(A)$ to denote, respectively, the symmetric and antisymmetric part of a square matrix $A \in \mathbb{R}^{d\times d}$, that is,

$$\mathrm{sym}(A) = \frac{1}{2}\left(A + A^{\mathrm{T}}\right), \qquad \mathrm{skew}(A) = \frac{1}{2}\left(A - A^{\mathrm{T}}\right). \tag{1}$$

2.2. Riemannian geometry. At a high level, a *manifold* $\mathcal{M}$ is defined by a topological space that is Hausdorff[1] and that is equipped with a set of overlapping local coordinate charts. These charts locally parametrize the space, and it is possible to transition smoothly from one chart to the other. The *tangent space* at a point $x \in \mathcal{M}$, denoted as $T_x\mathcal{M}$, can be defined as the linear space containing all the *tangent vectors* corresponding to the curves passing through x. We use the notation $v^\vee$ to denote the vector of coordinates of v in some basis for $T_x\mathcal{M}$. A *vector field* V assigns a tangent vector $v = V|_x$ to each x in $\mathcal{M}$ or a subset of it. We denote as $\mathcal{X}(\mathcal{M})$ the set of smooth vector fields on $\mathcal{M}$. Given $V, W \in \mathcal{X}(\mathcal{M})$, the *Lie bracket* of two vector fields is denoted as $[V, W] \in \mathcal{X}(\mathcal{M})$. Intuitively, $[V, W]$ represents the "derivative" of a field with respect to another and assumes a simple expression for the manifolds and fields considered in this paper (as we will see in subsection 2.7).

A Riemannian manifold $(\mathcal{M}, \langle\cdot,\cdot\rangle)$ is a manifold endowed with a *Riemannian metric*, a collection of inner products $\langle\cdot,\cdot\rangle_x$ over $T_x\mathcal{M}$ that varies smoothly with x. The metric is used to define the length of a curve $\gamma : \mathbb{R} \supset [a, b] \to \mathcal{M}$. A curve is a *geodesic* if the covariant derivative of its tangent is zero, i.e., $\nabla_{\dot\gamma}\dot\gamma \equiv 0$, where ∇ is the *Levi-Civita connection*. The *exponential map* $\exp_x : T_x\mathcal{M} \to \mathcal{M}$ maps each tangent vector v to the endpoint of the unit-speed geodesic starting at x with tangent v. The *logarithm map* $\log_x$ is the inverse of the $\exp_x$ and is defined (in general) only on a neighborhood of x. We use the shorthand notation $\mathrm{Log} = \log^\vee$ to denote the logarithm map expressed in local tangent space coordinates. For any point x and any curve $\gamma(t)$ in $\mathcal{M}$ sufficiently close together, the logarithm is related to the distance function by the relations

$$d\big(x, \gamma(t)\big) = \|\mathrm{Log}_x\big(\gamma(t)\big)\|, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}d^2\big(x, \gamma(t)\big) = -\mathrm{Log}_x\big(\gamma(t)\big)^{\mathrm{T}}\dot\gamma(t)^\vee. \tag{2}$$

Given the Levi-Civita connection, one can define an intrinsic notion of *curvature* of the space. There are different ways to rigorously capture this quantity. One of the simplest ones is the *sectional curvature* $K_{\sigma(v,w)}(x)$, which denotes the curvature of $\mathcal{M}$ at a point x when restricted to a subspace $\sigma(v, w) \subset T_x\mathcal{M}$ spanned by two linearly independent vectors $v, w \in T_x\mathcal{M}$. The exact definition of $K_{\sigma(v,w)}(x)$ is not needed in this paper; however, intuitively one can think of this quantity as a way to measure how fast two geodesics starting at x in the directions u and v either spread (negative curvature) or converge (positive curvature)

[1] A space is *Hausdorff* if, for any two distinct points $x, y \in \mathcal{M}$, $x \neq y$, there exists two disjoint open subsets, $U, V \subset \mathcal{M}$, such that $x \in U$ and $y \in V$.

with respect to similar geodesics in Euclidean space (which has zero curvature). In practice, knowing bounds on the curvature of the space allows one to derive convergence guarantees for optimization algorithms such as those described in [2, 27] and the Weiszfeld algorithm, which we will use in subsection 7.2.

2.3. Differentials, gradients, and Hessians. Let $f : \mathcal{M} \to \tilde{\mathcal{M}}$ be a map between manifolds. The map is said to be *proper* if $f^{-1}(\tilde{U})$ is compact in M for every compact subset $\tilde{U}$ in $\tilde{\mathcal{M}}$.

We define $\mathrm{D}f$ as the *differential* of the map, i.e., the linear operator (the Jacobian matrix, in local coordinates) that maps $T_x\mathcal{M}$ to $T_{f(x)}\tilde{\mathcal{M}}$ for any $x \in \mathcal{M}$ and satisfies, for any locally defined curve $\gamma(t) \in \mathcal{M}$, the expression

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(\gamma(t)\big) = \mathrm{D}f\big(\gamma(t)\big)[\dot{\gamma}]. \tag{3}$$

(To clarify, the left-hand side is the tangent of the curve $f\big(\gamma(t)\big)$, and the right-hand side is the differential $\mathrm{D}f$ computed at $\gamma(t)$ applied to the vector $\dot{\gamma}$.)

When f is a scalar function $f : \mathcal{M} \to \mathbb{R}$, the *Riemannian gradient* of f at $x \in \mathcal{M}$ is defined as the unique tangent vector $\operatorname{grad} f(x)$ such that, for all $v \in T_x\mathcal{M}$,

$$\langle \operatorname{grad} f(x), v \rangle = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\Big|_{t=0}, \tag{4}$$

where γ is a smooth curve passing through $\gamma(0) = x$ with tangent $\dot{\gamma}(0) = v$. Similarly, the *Riemannian Hessian* at x can be defined as the self-adjoint operator $\operatorname{Hess} f : T_x\mathcal{M} \to T_x\mathcal{M}$ satisfying

$$\langle v, \operatorname{Hess} f(x)[v] \rangle = \frac{\mathrm{d}^2}{\mathrm{d}t^2} f(\gamma(t))\Big|_{t=0}. \tag{5}$$

Intuitively, as in the Euclidean case, the gradient indicates the direction along which f increases the most, while the Hessian indicates its local quadratic behavior.

2.4. Lie groups. A *group* is a set G together with an operation $\circ : G \times G \to G$ that satisfy the usual four axioms of closure, associativity, identity, and inverse.

Given an element $g \in G$, the left (respectively, right) translation of G, L_g (respectively, R_g) is defined as $L_g : h \mapsto g \circ h$ (respectively, $R_g : h \mapsto h \circ g$); i.e., it is the mapping that multiplies each elements in G by a common element g on the left (respectively, right).

A *Lie group* is a group that is also a smooth manifold and for which the group and inverse operations are smooth mappings. In this case, the differentials $\mathrm{D}L_g$ and $\mathrm{D}R_g$ of left and right translations are well defined. A Riemannian metric on G is *left invariant* if, for all $g, h \in G$ and $v, w \in T_hG$,

$$\langle v, w \rangle_h = \langle \mathrm{D}L_g(h)[v], \mathrm{D}L_g(h)[w] \rangle_{L_g(h)}; \tag{6}$$

i.e., $L_g(h)$ is a local isometry for all h. A similar definition holds for a *right-invariant* metric. A metric is *bi-invariant* if it is both left and right invariant.

A field $V \in \mathcal{X}(G)$ is *left invariant* if $\mathrm{D}L_g[V] = V$ for all $g \in G$. In particular, this means that

$$V|_g = \mathrm{D}L_g[V|_e], \tag{7}$$

and we can identify any left-invariant field $V \in \mathcal{X}(G)$ with a vector in the tangent space at the identity, $v = V|_e \in T_eG$. It turns out that left-invariant vector fields are closed under the Lie bracket operation; i.e., if $V, W \in \mathcal{X}(G)$ are left invariant, then also $[V, W]$ is left invariant. The identification (7) then turns T_eG into the so-called *Lie algebra*, where the bracket operation is given by $[V|_e, W|_e] = [V, W]|_e$. For the manifolds considered in this paper, this will provide a simple expression for computing the Lie bracket of vector fields (see subsection 2.7).

In addition, if the metric is bi-invariant, the expression for the sectional curvature also assumes a particularly simple form:

$$K_{\sigma(v,w)}(g) = \frac{1}{4}\|[V, W]|_g\|^2, \tag{8}$$

where $V, W \in \mathcal{X}(G)$ are the left-invariant extensions of $v, w \in T_xG$.

2.5. Quotient spaces. A *group action* "$\cdot$" on a set $\mathcal{M}$ is a mapping $\cdot : G \times \mathcal{M} \to \mathcal{M}$ that satisfies the properties $g \cdot (h \cdot x) = (g \circ h) \cdot x$ and $e \cdot x = x$ for all $g, h \in G$, $x \in \mathcal{M}$. The action is said to be *free* if $g \cdot x = x$ for at least one $x \in \mathcal{M}$ implies that $g = e$. As an important particular case, if G is discrete and its action is free and proper (this is sometimes referred to as a *properly discontinuous* action [15]), then we have the property that different elements of g map the same neighborhood U of an arbitrary point x to disjoint sets, i.e.,

$$(g_1 \cdot U) \cap (g_2 \cdot U) \neq \emptyset \implies g_1 \neq g_2 \tag{9}$$

for any $g_1, g_2 \in G$. The group action induces an *equivalence relation* between the points in $\mathcal{M}$, and we say that x is equivalent to y, i.e., $x \sim y$, if there exist $g \in G$ such that $g \cdot x = y$. We denote all the elements equivalent to $x \in M$ as the *equivalence class* (also called *orbit*) $[x]$. The *quotient space* $\mathcal{M}/G$ is the space of all equivalence classes. The *canonical projection* $\pi : \mathcal{M} \to \mathcal{M}/G$ maps each point $x \in \mathcal{M}$ to $[x]$.

2.6. Riemannian submersions. A map $f : \mathcal{M} \to \tilde{\mathcal{M}}$, where $\dim(\mathcal{M}) > \dim(\tilde{\mathcal{M}})$, is said to be a *submersion* if $\mathrm{D}f$ is injective (i.e., as a matrix, it has full rank) on the entire domain $\mathcal{M}$. An example of submersion is the canonical projection π from a manifold $\mathcal{M}$ to the quotient $\tilde{\mathcal{M}} = \mathcal{M}/G$ when the action of the group G is proper and free.

If f is a submersion, then the differential of f, $\mathrm{D}f : T_x\mathcal{M} \to T_{f(x)}\mathcal{M}$ has full rank, and $T_x\mathcal{M}$ admits the following orthogonal decomposition:

$$T_x\mathcal{M} = T_{Vx}\mathcal{M} \oplus T_{Hx}\mathcal{M}, \tag{10}$$

where the *vertical space* $T_{Vx}\mathcal{M}$ is equal to $\ker(\mathrm{D}f)$, the kernel of $\mathrm{D}f$, and the *horizontal space* $T_{Hx}\mathcal{M}$ is its orthogonal complement $\ker^{\perp}(\mathrm{D}f)$. Intuitively, for the canonical projection π, $T_{Vx}\mathcal{M}$ contains the vectors tangent to the equivalence class $[x] = \pi(x)$, while $T_{Hx}\mathcal{M}$ contains vectors pointing between different classes. We denote the orthogonal projection of a vector $v \in T_x\mathcal{M}$ on the horizontal and vertical spaces as, respectively, $\mathcal{H}v$ and $\mathcal{V}v$. Note that, from the properties of orthogonal projections, it follows that, for horizontal vectors $v_H \in T_{Hx}\mathcal{M}$,

$$\langle v_H, v\rangle = \langle v_H, \mathcal{H}v\rangle, \tag{11}$$

with an analogous expression for vertical vectors. With the decomposition (10), we can associate any vector $v \in T_{f(x)}\tilde{\mathcal{M}}$ to a unique vector $\bar{v} \in T_{Hx}\mathcal{M}$, called the *horizontal lift* of v.

An important fact about submersions is that we can relate the Riemannian connections of the ambient space to that in the submerged manifold (see [19, 20]). This, in turn, gives relations between the respective geodesics and curvatures.

Proposition 2.1. *Let $\gamma(t) : \mathbb{R} \to \mathcal{M}$ be a geodesic curve such that $\dot{\gamma}(t) \in T_{H\gamma}\mathcal{M}$ for all t. Then $\tilde{\gamma} = f(\gamma)$ is a geodesic curve in $\tilde{\mathcal{M}}$.*

Proof. Denote as $\nabla^{\mathcal{M}}$ and $\nabla^{\tilde{\mathcal{M}}}$ the Levi-Civita connections for $\mathcal{M}$ and $\tilde{\mathcal{M}}$, respectively. From [19, 20], we have that $\mathcal{H}(\nabla^{\mathcal{M}}_X Y)$ is the horizontal lift of $\nabla^{\tilde{\mathcal{M}}}_{\tilde{X}}\tilde{Y}$, where $\tilde{X}, \tilde{Y}$ are vector fields on $\tilde{\mathcal{M}}$ and V, W are their horizontal lifts on $\mathcal{M}$. The defining property of the geodesic $\gamma(t)$ is that $\nabla^{\mathcal{M}}_{\dot{\gamma}}\dot{\gamma} = 0$. However, since $\dot{\gamma}(t)$ is always horizontal, and given the isometry between $T_{H\gamma}\mathcal{M}$ and $T_{\dot{\gamma}}\tilde{\mathcal{M}}$, we have that $\dot{\gamma}$ is the horizontal lift of $\dot{\tilde{\gamma}}$. Then we have $\nabla^{\tilde{\mathcal{M}}}_{\dot{\tilde{\gamma}}}\dot{\tilde{\gamma}} = 0$ because $\mathcal{H}\nabla^{\mathcal{M}}_{\dot{\gamma}}\dot{\gamma} = \nabla^{\mathcal{M}}_{\dot{\gamma}}\dot{\gamma} = 0$. Hence, $\tilde{\gamma}$ is a geodesic in $\tilde{\mathcal{M}}$. ∎

Let $K_{\sigma(v,w)}$ and $\tilde{K}_{\sigma(\tilde{v},\tilde{w})}$ denote the sectional curvatures in $\mathcal{M}$ and $\mathcal{M}/G$. Then the two are related by the formula [21]

$$\tilde{K}_{\sigma(\tilde{v},\tilde{w})}(x) = K_{\sigma(v,w)}(x) + \frac{3}{4}\|\mathcal{V}[V,W]|_x\|^2, \tag{12}$$

where $v, w \in T_{Hx}\mathcal{M}$ are the horizontal lifts of $\tilde{v}, \tilde{w} \in T_{[x]}\mathcal{M}$ and $V, W \in \mathcal{X}$ are any smooth horizontal extensions of v, w to a neighborhood of x. For the adept reader, we remark that the quantity $\mathcal{V}[V, W]$ can be shown to be tensorial for V, W horizontal [21]; hence, it depends only on the point-wise values v, w and not on their particular extensions V, W.

2.7. The Lie groups of rotations $SO(3)$ and $SO(3) \times SO(3)$. In this paper, we will heavily use the space of 3-D rotations $SO(3) = \{R \in \mathbb{R}^{3\times 3} : R^{\mathrm{T}}R = I, \det(R) = 1\}$ and, to a lesser extent, the space of rigid body transformations $SE(3) = SO(3) \ltimes \mathbb{R}^3$. We will also use $SO(3) \times SO(3)$, the Cartesian product of $SO(3)$ with itself. The spaces $SO(3)$ and $SE(3)$ are *Lie groups*. For $SO(3)$, the group operation corresponds to matrix multiplication with I as the identity element. For $SE(3)$, the group operation is the semidirect product given by $(R_1, T_1) \circ (R_2, T_2) = (R_1R_2, R_1T_2 + T_1)$. The space $SO(3) \times SO(3)$ can also be interpreted as a Lie group with the group operation acting component-wise; i.e., for $(R_1, R_2), (S_1, S_2) \in SO(3) \times SO(3)$, we have $(R_1, R_2) \circ (S_1, S_2) = (R_1S_1, R_2S_2)$. The identity element of this group is simply (I, I). For convenience, one can also represent $SO(3) \times SO(3)$ as a subset of $\mathbb{R}^{6\times 6}$ with the embedding

$$(R_1, R_2) \mapsto \operatorname{diag}(R_1, R_2). \tag{13}$$

The group operation is then the same as matrix multiplication.

The tangent space at $R \in SO(3)$ is $T_RSO(3) = \{RV : V \in \mathfrak{so}(n)\}$. The Lie algebra of $SO(3)$ is then $T_ISO(3) = \mathfrak{so}(3)$, with the bracket operation in the Lie algebra given by the matrix commutator (this comes from the fact that $SO(3)$ is a subset of the space of 3-by-3 matrices; see, e.g., [21, Lemma 70, page 378]):

$$[V, W] = VW - WV, \qquad V, W \in \mathfrak{so}(3). \tag{14}$$

For $SO(3) \times SO(3)$, the Lie algebra is simply $\mathfrak{so}(3) \times \mathfrak{so}(3)$, and the bracket is defined component-wise, i.e.,

$$[(V_1, V_2), (W_1, W_2)] = ([V_1, W_1], [V_2, W_2]), \qquad (V_1, V_2), (W_1, W_2) \in \mathfrak{so}(3)^2. \tag{15}$$

The easiest way to see this is to use [21, Lemma 70, page 378] on the embedding (13).

We can identify a tangent vector $v \in T_R SO(3)$ with a vector of local coordinates $w \in \mathbb{R}^3$ using the *hat* $(\cdot)^\wedge$ and *vee* $(\cdot)^\vee$ operators, given by the relations

$$v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} \in \mathbb{R}^3 \underset{(\cdot)^\vee}{\overset{(\cdot)^\wedge}{\rightleftarrows}} V = R \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix} \in T_R SO(3). \tag{16}$$

This notation can be easily extended to $SO(3) \times SO(3)$ with local coordinates $w \in \mathbb{R}^6$ given by

$$v = \operatorname{stack}(v_1, v_2) = \operatorname{stack}(V_1^\vee, V_2^\vee). \tag{17}$$

From the identifications (7) and (16), we can associate a left-invariant vector field V to any coordinate vector $v \in \mathbb{R}^3$. In fact, one can check with a direct computation that the bracket operation (14) becomes, in coordinate vectors,

$$[(v)^\wedge, (w)^\wedge]^\vee = [v]_\times w. \tag{18}$$

Note that, from the fact that ${[v]_\times}^2 = v^{\mathrm{T}} v - \|v\|^2 I$, we have

$$\|[v]_\times w\|^2 = \|v\|^2 \|w\|^2 - (v^{\mathrm{T}} w)^2; \tag{19}$$

hence,

$$0 \leq \|[v]_\times w\|^2 \leq \|v\|^2 \|w\|^2. \tag{20}$$

The standard metric for $SO(3)$, with $V, W \in T_R SO(3)$, is given by

$$\langle V, W \rangle = \frac{1}{2} \operatorname{tr}(V^{\mathrm{T}} W) = (V^\vee)^{\mathrm{T}} W^\vee, \tag{21}$$

where we used again the identification (16). A similar expression holds for $SO(3) \times SO(3)$. It can be easily shown that this metric is bi-invariant.

The exponential map and logarithm maps for $SO(3)$ are given by

$$\exp_R(V) = R \operatorname{expm}(R^{\mathrm{T}} V) \tag{22}$$

$$\log_R(S) = R \operatorname{logm}(R^{\mathrm{T}} S) \tag{23}$$

where $R, S \in SO(3)$, $V \in T_{SO(3)} R$, expm is the matrix exponential defined as

$$\operatorname{expm}(A) = \sum_{i=1}^{\infty} \frac{1}{k!} A^k \tag{24}$$

and logm is its inverse (which is well defined in a neighborhood of the identity matrix). These last two maps applied on skew-symmetric matrices and rotations as in (22) and (23) can be computed in closed form using Rodrigues' formula (see [17]).

We denote as $R_x(\theta)$, $R_y(\theta)$, and $R_z(\theta)$ the rotations around the x, y, and z axes, respectively, with angle $\theta \in [-\pi, \pi)$, and as e_z the unit vector aligned with the z axis.

With the standard metric, we can substitute (18) into (8); from (19) and the fact that v, w must be orthonormal, this leads to the well-known fact that, for $SO(3)$,

$$K^{SO(3)}_{\sigma(v,w)}(x) = \frac{1}{4}\|[v]_\times w\|^2 = \frac{1}{4}; \tag{25}$$

i.e., $SO(3)$ has constant positive curvature.

Similarly, for $SO(3) \times SO(3)$,

$$K^{SO(3)}_{\sigma(v,w)}(x) = \frac{1}{4}\big(\|[v_1]_\times w_1\|^2 + \|[v_2]_\times w_2\|^2\big). \tag{26}$$

Hence, from (20) and the fact that $\|v_1\|, \|v_2\|, \|w_1\|\|w_2\| \leq 1$ (since $\|v\| = \|w\| = 1$), we obtain that

$$0 \leq K^{SO(3)\times SO(3)}_{\sigma(v,w)}(x) \leq \frac{1}{2}; \tag{27}$$

i.e., $SO(3) \times SO(3)$ has nonnegative curvature.

3. Derivation of the essential matrix. As customary, we model the pose of the ith camera, $i \in \{1, 2\}$, as $g_i = (R'_i, T'_i) \in SE(3)$, where g_i represents the transformation from camera to world coordinates. Given an image x_i in homogeneous coordinates and the corresponding depth λ_i, the 3-D point in world coordinates is given by

$$X = \lambda_i R'_i x_i + T'_i. \tag{28}$$

Note that a change of world coordinates represented by $g = (R_0, T_0)$, i.e., $X \mapsto R_0 X + T_0$, induces a transformation of the camera representation equivalent to multiplying g_i by g on the left, i.e., $(R'_i, T'_i) \mapsto (R_0 R'_i, R_0 T'_i + T_0)$.

We now derive the essential matrix from an epipolar configuration, i.e., two camera poses (R'_i, T'_i), and the two images (x_1, x_2), of a same 3-D point X. We follow a general approach [4] as opposed to the traditional one, which uses one camera as the global reference frame. From (28) and using the properties $[a]_\times a = 0$ and $b^{\mathrm{T}}[a]_\times b = 0$ for all $a, b \in \mathbb{R}^3$, we have

$$\lambda_1 R'_1 x_1 + T'_1 = \lambda_2 R'_2 x_2 + T'_2 \tag{29}$$

$$\lambda_1 R'_1 x_1 = \lambda_2 R'_2 x_2 + (T'_2 - T'_1) \tag{30}$$

$$\lambda_1 [T'_2 - T'_1]_\times R'_1 x_1 = \lambda_2 [T'_2 - T'_1]_\times R'_2 x_2 \tag{31}$$

$$x_1^{\mathrm{T}} {R'_1}^{\mathrm{T}} [T'_2 - T'_1]_\times R'_2 x_2 = 0. \tag{32}$$

The essential matrix is then defined as

$$E = {R'_1}^{\mathrm{T}} [T'_2 - T'_1]_\times R'_2. \tag{33}$$

Algebraic and geometric interpretations of the essential matrix. It is possible to interpret an essential matrix $E \in \mathbb{R}^{3\times 3}$ in two ways: as a matrix of coefficients in the epipolar constraint (32) or as a way to encode the relative pose of two cameras. Due to the cancellations performed during the derivation of (32), these two views are similar but not exactly equivalent. If we limit ourselves to the first interpretation, we cannot resolve the twisted-pair ambiguity, where four different epipolar configurations produce the same essential matrix. However, to consider the second interpretation, we need to use additional information (pairs of corresponding image points (x_1, x_2)) to solve the same ambiguity. In this paper, we consider both interpretations. The two interpretations will, respectively, give rise to what we call the *normalized essential manifold* (sections 4 and 6) and the *signed normalized essential manifold* (section 5). We will show that the difference between the two geometries is exactly given by the twisted-pair ambiguity, that the latter is a Riemannian covering of the former, and that we can pass from one to the other using data from corresponding image points.

4. The normalized essential space.

In this section, we define a canonical decomposition of the essential matrix in terms of two rotations by choosing a global reference frame aligned with the baseline between the two cameras. Then we define the *normalized essential space* and analyze its structure as a quotient space and its relation with epipolar configurations, the twisted-pair ambiguity, and transformations of image vectors.

4.1. The normalized canonical decomposition.

Since (32) is a homogeneous equation, we cannot determine the scale and sign of E from image data alone. Also, while E does not depend on the choice of global reference frame, this is not true for its decomposition (33). To remove most of the degrees of freedom, we define $Q = (R_1, R_2) \in SO(3) \times SO(3)$ and use the following.

Proposition 4.1. *Any essential matrix E admits, up to scale, the following* normalized canonical decomposition*:*

$$E(Q) = R_1^{\mathrm{T}}[e_z]_\times R_2. \tag{34}$$

Proof. Starting from (33), choose a global scale such that $\|T_2' - T_1'\| = 1$ and let $R_0 \in SO(3)$ be such that $R_0(T_2' - T_1') = e_z$. There are infinite candidates for such rotation (we pick one using Householder transformations). Then, by applying the transformation $g_0 = (R_0, 0)$ and using the property $R[a]_\times R^{\mathrm{T}} = [Ra]_\times$ for all $R \in SO(3)$, we have

$$E = (R_0 R_1')^{\mathrm{T}}[R_0(T_2' - T_1')]_\times R_0 R_2', \tag{35}$$

which is of the form (34) with $R_i = R_0 R_i'$, $i = 1, 2$. ∎

Intuitively, the change of world coordinates performed in the proof above aligns the vector $T_2' - T_1'$ with the z-axis. In this way, the translation direction is known, and we are left with only the information about the two rotations.

Remark 4.2. Notice that $[e_z]_\times R_z\left(\frac{\pi}{2}\right) = P_z = \mathrm{diag}(1, 1, 0)$. Hence, $E = R_1^{\mathrm{T}} P_z (R_z\left(\frac{\pi}{2}\right) R_2)$ is a valid SVD of E with factors R_1, P_z, and $R_z\left(\frac{\pi}{2}\right) R_2$.

The value of Remark 4.2 is twofold. First, it provides a practical way to compute the decomposition (34). Second, it relates our representation with the one of [24], giving a geometric meaning to the SVD of E.

Note that (34) defines a map $Q \mapsto E(Q)$ from $SO(3) \times SO(3)$ to $\mathbb{R}^{3\times 3}$. We define the *normalized essential space* $\mathcal{M}_E$ as the image of such map. Since, according to Proposition 4.1, (34) is surjective, $\mathcal{M}_E$ corresponds to the space of all the essential matrices.

4.2. Ambiguities of the canonical form. While the map (34) is surjective, it cannot be also injective because it is known that the space of essential matrices is five-dimensional, while $SO(3) \times SO(3)$ is six-dimensional. The extra degree of freedom corresponds to a rotation of the global reference frame around the baseline (i.e., to a particular choice of R_0 in the proof of Proposition 4.1). However, it turns out that this is not the only ambiguity. To be more precise, consider any two points $Q_a, Q_b \in SO(3) \times SO(3)$, which, through (34), correspond to the essential matrices E_a, E_b. We define an equivalence relation "$\sim$" between points in $SO(3) \times SO(3)$ as

$$Q_a \sim Q_b \iff E_a = E_b, \tag{36}$$

where, again, equality is intended up to scale (since E_a and E_b are normalized, this reduces to "up to a sign flip").

Proposition 4.3. *Define the groups*

$$H_z = \{\big(R_z(\theta), R_z(\theta)\big) : \theta \in [-\pi, \pi)\}, \tag{37}$$

$$H_\pi = \big\{(I, I), \big(R_y(\pi), R_y(\pi)\big) \big(I, R_z(\pi)\big), \big(R_y(\pi), R_x(\pi)\big)\big\} \tag{38}$$

acting on the left on $SO(3) \times SO(3)$ by simple component-wise left multiplication. Then the equivalence class of a point Q with respect to "$\sim$" is exactly given by[2]

$$[[Q]] = \{S_\pi S_z Q : S_\pi \in H_\pi, S_z \in H_z\}. \tag{39}$$

The proof involves first showing that H_z and H_π are subgroup of $SO(3) \times SO(3)$, and then showing, with an exhaustive case-by-case argument, that the only matrices satisfying (39) are those in the equivalence class $[[Q]]$.[3] Since the proof is quite long and mostly mechanical, it has been moved to Appendix A. In the following, we will use $S_z = (S_{z1}, S_{z2})$ and $S_\pi = (S_{\pi 1}, S_{\pi 2})$ to denote points in H_z and in H_π, respectively.

Remark 4.4. One can easily verify by direct computation that the group H_π can be decomposed as $H_\pi = H_{y\pi} \times H_{z\pi}$, where

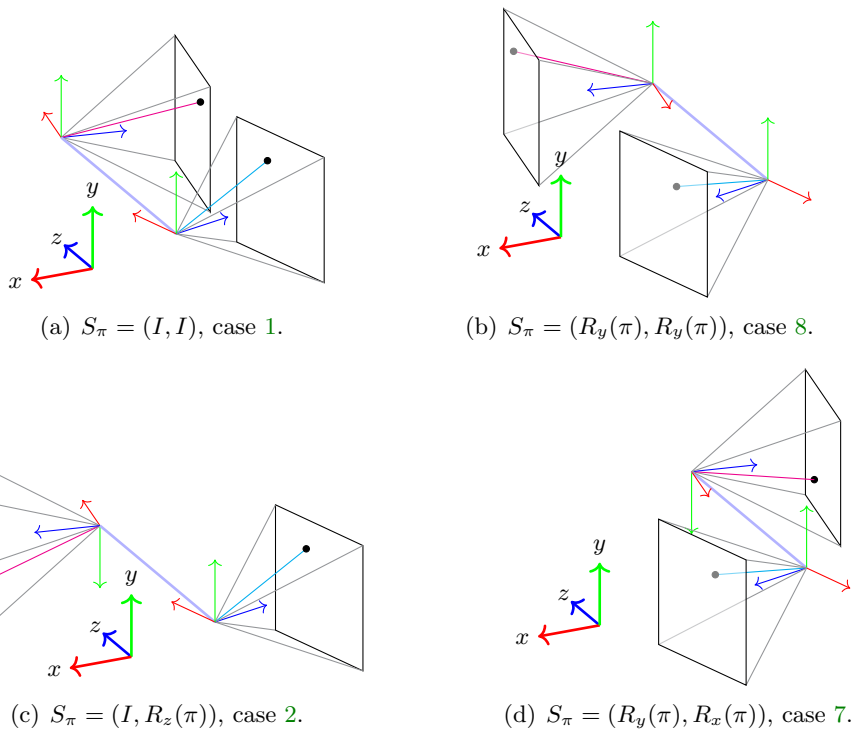
$$H_{y\pi} = \big\{(I, I), \big(R_y(\pi), R_y(\pi)\big)\big\} \tag{40}$$

$$H_{z\pi} = \big\{\big(I, I\big), \big(I, R_z(\pi)\big)\big\} \tag{41}$$

which are both isomorphic to the cyclic group $Z_2 = \{1, -1\}$. This shows that H_π is isomorphic to the *Klein four-group*, also known as the symmetry group of the 2-D plane produced by 180° rotations and reflection across one of the axes.

[2]The definition of H_π and the order of S_π, S_z are different with respect to [28]. This does not change the equivalence class $[[Q]]$, but it facilitate the derivations for the (unsigned) essential manifold in section 6.

[3]The use of double brackets $[[\cdot]]$ is due to the fact that we consider two groups, H_z and H_π, instead of one. This will become clear after section 6.



(a) $S_\pi = (I, I)$, case 1. (b) $S_\pi = (R_y(\pi), R_y(\pi))$, case 8.

(c) $S_\pi = (I, R_z(\pi))$, case 2. (d) $S_\pi = (R_y(\pi), R_x(\pi))$, case 7.

Figure 1. *Diagram depicting the geometric twisted-pair ambiguity given by the four elements of H_π. The case numbers refer to the choices of signs in the proof of Proposition 4.3 in Appendix A.*

Remark 4.5. The result of Proposition 4.3 is equivalent to the traditional proof that each essential matrix can be factorized in four epipolar configurations (e.g., [17]), where the poses are parametrized as (R_2, T_2), i.e., the rotation and normalized translation of the second camera with respect to the first. An alternative proof to Proposition 4.3 could then be formulated by first obtaining (R_2, T_2) from Q, use the traditional result, and then return to the quotient representation. The proof in Appendix A shows that this is not necessary and that one can work directly with the quotient representation. Although the proof is algebraic in nature, it has also some geometric interpretation (see Figure 1 and subsection 4.3 below).

Intuitively, one can visualize $[[Q]]$ has having four components, where each one of the components is isomorphic to $SO(2)$ (i.e., the circle). In view of Proposition 4.3, the space $\mathcal{M}_E$ can be identified with the quotient space

$$\mathcal{M}_E = \big(SO(3) \times SO(3)\big)/(H_\pi \times H_z), \tag{42}$$

where the actions of H_π and H_z are defined above.

Since $SO(3) \times SO(3)$ has dimension 6 and H_z has dimension 1, we get the well-known fact that the normalized essential space has dimension 5 (being discrete, H_π does not change the intrinsic dimension of the space).

Remark 4.6. The normalized essential *space* $\mathcal{M}_E$ is actually a Riemannian *manifold*. We will provide this characterization in section 6 after the analysis of the simpler quotient space $\big(SO(3) \times SO(3)\big)/H_z$ in section 5.

4.3. Geometric interpretation with image vectors. Using the geometric interpretation given by the proof of Proposition 4.1, we now show that also the epipolar constraint $x_1^{\mathrm{T}} E x_2 = 0$ has a geometrical interpretation in terms of our parametrization. Given an essential matrix $E = R_1^{\mathrm{T}}[e_z]_\times R_2$, from Proposition 4.3 and the equivalence $[e_z]_\times = P_z^{\mathrm{T}} R_z\big(\frac{\pi}{2}\big)\, P_z$, we have

$$x_1^{\mathrm{T}} E x_2 = (P_z S_{\pi 1} S_z R_1 x_1)^{\mathrm{T}} R_z\left(\frac{\pi}{2}\right)(P_z S_{\pi 2} S_z R_2 x_2) = 0. \tag{43}$$

This can be interpreted as the following procedure:

- Take the images x_i and rotate them as $R_i x_i$, $i = 1, 2$. This is equivalent to expressing in global coordinates the vectors corresponding to the images and centering them at the origin. Notice that, by construction, the transformed vectors and the z-axis e_z all lie in the same plane passing through the origin.
- Apply the action of an element of H_z; i.e., rotate the two vectors around the z-axis by an arbitrary amount. This is equivalent to a rotation around the baseline and does not change the coplanarity condition.
- Apply the action of an element of $S_\pi = H_\pi$ (see Figure 1). If $S_\pi = (I, I)$, no changes are made. Otherwise, the direction of the first, second, or both cameras is reversed (i.e., the rotated cameras pass from front-facing to rear-facing). Note that the coplanarity condition of the transformed vectors with e_z is preserved.
- Project the transformed vectors onto the xy-plane. In practice, this sets the last coordinate to zero. Before the projection, the vectors belonged to the same plane, and this plane contained e_z; after the projection, the vectors will have the same direction (but generally different lengths), and they will be orthogonal to e_z.
- Rotate one of the projected vectors by $R_z\big(\frac{\pi}{2}\big)$, e.g., $R_z\big(\frac{\pi}{2}\big)\,(P_z R_2 x_2)$. Since the vectors were collinear, they are now orthogonal, and the inner product is zero.

In our context, this interpretation of (43) shows that the action of H_π corresponds exactly to the well-known twisted-pair ambiguity in the decomposition of the essential matrix. In fact, the four cases of Figure 1 correspond to the four valid cases in the proof of Proposition 4.3 in Appendix A.

5. The signed normalized essential manifold. In this section, we review how the chirality constraint can be used to resolve the twisted-pair ambiguity (i.e., to choose an element of the group H_π). This can be used to "unfold" the quotient structure of the normalized essential space into what we call the *signed normalized essential space*. Intuitively, this is the space of geometrically distinct epipolar configurations (e.g., the four configurations in Figure 1 correspond to four different points in the unfolded space). We show that this space is a manifold and that a metric and the corresponding geodesics can be induced from $SO(3) \times SO(3)$. Finally, we give a Newton-based optimization algorithm for computing the logarithm map and the Riemannian distance.

5.1. Depth triangulation. We can use the simple geometrical interpretation of the canonical form to estimate the depths of the 3-D points and enforce the chirality constraint, i.e., the fact that all these points need to be in front of both cameras.

From the discussion in subsection 4.2, we have $T_2' - T_1' = e_z$ in the canonical form. Therefore, taking into account H_z and H_π and assuming noiseless image points, (30) becomes

$$\lambda_1 S_{z1} S_{\pi 1} R_1 x_1 = \lambda_2 S_{z1} S_{\pi 2} R_2 x_2 + e_z. \tag{44}$$

Note that $e_z = S_{z1} e_z = S_{z2} e_z$; hence, we can cancel S_z from (44). We then have the following.

Proposition 5.1. *There is only one choice of S_π for which the solution of (44) is positive, i.e., $\lambda_1, \lambda_2 > 0$.*

The proof is similar to the one traditionally used to solve the twisted-pair ambiguity [17].

Proof. Let r_{13}^{T} and r_{23}^{T} denote the last row of R_1 and R_2, respectively. Let $(\lambda_{j1}^*, \lambda_{j2}^*)$, $j \in \{1, \dots, 4\}$ be the solutions to (44) for each choice of S_π, in the order given by (38). One can verify that the last three solutions are related to the first one by

$$(\lambda_{21}^*, \lambda_{22}^*) = (-\lambda_{11}^*, -\lambda_{12}^*) \tag{45}$$

$$(\lambda_{31}^*, \lambda_{32}^*) = \left(\frac{\lambda_{11}^*}{a}, -\frac{\lambda_{12}^*}{a}\right) \tag{46}$$

$$(\lambda_{41}^*, \lambda_{42}^*) = \left(-\frac{\lambda_{11}^*}{a}, \frac{\lambda_{12}^*}{a}\right) \tag{47}$$

where $a = \lambda_{11}^* r_{13}^{\mathrm{T}} x_1 + \lambda_{12}^* r_{23}^{\mathrm{T}} x_2$. Notice that, independently from the sign of λ_{11}^*, λ_{12}^* and a, the four solutions always cover all the possible sign combinations $(+,+)$, $(+,-)$, $(-,+)$, $(-,-)$. Hence, only one solution has both $\lambda_{j1}^* > 0$ and $\lambda_{j2}^* > 0$. ∎

As a concrete example, if one imagines intersecting the lines given by the cyan and magenta image vectors in Figure 1, the resulting triangulated point is in front of both cameras only in configuration (a).

5.2. The signed normalized essential space. In our context, Proposition 5.1 allows us to pick one of four components in the equivalent class $[[Q]]$. Looking at the definition of $[[Q]]$ in (39), this means that we can dispense with the group H_π and consider a new quotient space using H_z alone, which we call the *signed essential space*. Just for fun, we use the symbol $\mathcal{M}_{\exists}$ (because it differs from "$\mathcal{M}_E$" by a 180-degree rotation). Formally, we have

$$\mathcal{M}_{\exists} = \big(SO(3) \times SO(3)\big)/H_z, \tag{48}$$

and the equivalent class $[Q] \in \mathcal{M}_{\exists}$ corresponding to a point $Q \in SO(3) \times SO(3)$ is given by

$$[Q] = \{S_z Q : S_z \in H_z\}. \tag{49}$$

Intuitively, $[Q]$ contains all the epipolar configurations that are geometrically equivalent, i.e., that differ only by a rotation around the baseline (after aligning the baseline with e_z). Notice that $[Q]$ has dimension 1, the same as H_z.

5.3. A Riemannian quotient manifold structure. In general, a quotient space of a Riemannian manifold is not a Riemannian manifold itself. This is because the quotient might not be Hausdorff and hence fail to be a manifold at all; alternatively, the choice of a metric might not be obvious because it might depend on the choice of the representative in the equivalence class. However, the action of H_z has some "nice" properties that make $\mathcal{M}_{\mathcal{I}}$ into a Riemannian manifold with a natural choice for the metric. In order to show this, the first step is the following.

Proposition 5.2. *The canonical projection $\pi_{\mathcal{M}_{\mathcal{I}}} : SO(3) \times SO(3) \to \mathcal{M}_{\mathcal{I}}$ is a submersion, and $\mathcal{M}_{\mathcal{I}}$ is a manifold.*

Proof. The action of S_z is proper (which is automatic since $SO(3)$ is a compact manifold) and free (if $R_z R = R$, then multiply on the right by R^{T}, and this implies $R_z = I$). Then Theorem 9.16 from [15] proves the desired claim. ∎

Horizontal and vertical spaces. The fact that $\pi_{\mathcal{M}_{\mathcal{I}}}$ is a submersion implies that $T_Q\big(SO(3) \times SO(3)\big)$ admits an orthogonal decomposition in horizontal and vertical spaces (see subsection 2.5):

$$T_Q\big(SO(3) \times SO(3)\big) = T_{VQ}\big(SO(3) \times SO(3)\big) \oplus T_{HQ}\big(SO(3) \times SO(3)\big). \tag{50}$$

To obtain a concrete expression for the vertical space, we can differentiate a curve contained in an equivalence class $[Q]$. Let $\gamma(t) = (R_z(t)Q_1, R_z(t)Q_2)$ be the curve passing through $\gamma(0) = Q = (Q_1, Q_2)$ (by definition, $\gamma(t) \in [Q]$ for all $t \in \mathbb{R}$). Defining $v_V = \dot{\gamma}(0) \in T_{VQ}\big(SO(3) \times SO(3)\big)$, we have

$$v_V = ([e_z]_\times Q_1, [e_z]_\times Q_2) = \big(Q_1[Q_1^{\mathrm{T}} e_z]_\times, Q_2[Q_2^{\mathrm{T}} e_z]_\times\big). \tag{51}$$

The last equality in (51) is used to express the tangent vectors in the form of the right-hand side of (16), so that we can also write v_V as a coordinate vector:

$$(v_V)^\vee = \begin{bmatrix} v_{V1} \\ v_{V2} \end{bmatrix} = \begin{bmatrix} Q_1^{\mathrm{T}} e_z \\ Q_2^{\mathrm{T}} e_z \end{bmatrix}. \tag{52}$$

Since $[Q]$ has dimension 1, we have $T_{VQ}\big(SO(3) \times SO(3)\big) = \operatorname{span}(v_V)$.

Then, by definition, the horizontal space at Q includes all vectors v_H such that $v_H \perp v_V$, i.e.,

$$0 = \langle v_V, v_H \rangle = \big(Q_1^{\mathrm{T}} e_z\big)^{\mathrm{T}} v_{H1} + \big(Q_2^{\mathrm{T}} e_z\big)^{\mathrm{T}} v_{H2} = e_z^{\mathrm{T}}(Q_1 v_{H1} + Q_2 v_{H2}), \tag{53}$$

where $(v_H)^\vee = \operatorname{stack}(v_{H1}, v_{H2})$. We can take (53) as the condition defining horizontal vectors at Q. Given a vector $v \in T_Q SO(3) \times SO(3)$, let

$$p_Q(v) = e_z^{\mathrm{T}}(Q_1 v_1 + Q_2 v_2). \tag{54}$$

Using coordinate vectors, the orthogonal projection of v onto $T_{HQ}\big(SO(3) \times SO(3)\big)$ is then

$$(\mathcal{H}v)^\vee = v^\vee - \frac{p_Q(v)}{2} \begin{bmatrix} Q_1^{\mathrm{T}} e_z \\ Q_2^{\mathrm{T}} e_z \end{bmatrix}. \tag{55}$$

We will not explicitly use (55) in our theoretical derivations below, but we include this expression nonetheless because it is necessary when implementing the gradient and Hessian operators (see subsection 7.1).

Choice of metric. Since we know that $\mathcal{M}_{\mathcal{I}}$ is a manifold, the next step is to choose a Riemannian metric. We use the induced metric on $\mathcal{M}_{\mathcal{I}}$ obtained by using horizontal lifts to "borrow" the metric from $SO(3) \times SO(3)$; that is, we define

$$\langle \tilde{u}, \tilde{v} \rangle_{[Q]} = \langle u, v \rangle_Q, \tag{56}$$

where $v, w \in T_Q\big(SO(3) \times SO(3)$ are the horizontal lifts of $\tilde{v}, \tilde{w} \in T_{[Q]}\mathcal{M}_{\mathcal{I}}$. The following proposition shows that (56) is well posed.

Proposition 5.3. *The metric $\langle , \rangle_{[Q]}$ defined in (56) does not depend on the choice of the representative Q.*

The proof shows and relies on the fact that H_z is a group of global isometries for $SO(3) \times SO(3)$.

Proof. We first need to understand how to relate the horizontal lifts $v_a \in T_{Q_a}\big(SO(3) \times SO(3)\big), v_b \in T_{Q_b}\big(SO(3) \times SO(3)\big)$, $Q_a \sim Q_b$ of a same vector $\tilde{v}$. Let $S_z \in H_z$ such that $Q_b = S_z Q_a$. From the definition of the quotient space, we have

$$\pi_{\mathcal{M}_{\mathcal{I}}}(Q_a) = \pi_{\mathcal{M}_{\mathcal{I}}}(S_z Q_a). \tag{57}$$

Define arbitrary smooth curves $Q_a(t)$, $S_z(t)$ with $\dot{Q}_a = v_a$ horizontal. Differentiating (57), we have

$$\mathrm{D}\pi_{\mathcal{M}_{\mathcal{I}}}(Q_a)[\dot{Q}_a] = \mathrm{D}\pi_{\mathcal{M}_{\mathcal{I}}}(S_z Q_a)[\dot{S}_z Q_a + S_z \dot{Q}_a]. \tag{58}$$

However, notice that

$$\begin{aligned} \left(\dot{S}_z Q_a\right)^\vee &= \left(\dot{S}_{z1} Q_{a1}, \dot{S}_{z2} Q_{a2}\right)^\vee = (S_{z1}[e_z]_\times Q_{a1}, S_{z2}[e_z]_\times Q_{a2}) \\ &= (S_{z1} Q_{a1} \left(Q_{a1}^{\mathrm{T}} S_{z1}^{\mathrm{T}} e_z\right)^\wedge, S_{z2} Q_{a2} \left(Q_{a2}^{\mathrm{T}} S_{z2}^{\mathrm{T}} e_z\right)^\wedge) = \begin{bmatrix} Q_{a1}^{\mathrm{T}} e_z \\ Q_{a2}^{\mathrm{T}} e_z \end{bmatrix}, \end{aligned} \tag{59}$$

where we used the fact that $S_{z1}^{\mathrm{T}} e_z = S_{z2}^{\mathrm{T}} e_z = e_z$. Comparing (59) with (52), we see that $\dot{S}_z Q_a$ is vertical, and hence $\mathrm{D}\pi_{\mathcal{M}_{\mathcal{I}}}(S_z Q_a)[\dot{S}_z Q_a] = 0$, and this term can be canceled from (58). Moreover, since v_a is horizontal, using (53) we have

$$\begin{aligned} 0 &= \left(Q_{a1}^{\mathrm{T}} e_z\right)^{\mathrm{T}} v_{a1} + \left(Q_{a2}^{\mathrm{T}} e_z\right)^{\mathrm{T}} v_{a2} = \left(S_{z1} Q_{a1}^{\mathrm{T}} e_z\right)^{\mathrm{T}} S_{z1} v_{a1} + \left(S_{z2} Q_{a2}^{\mathrm{T}} e_z\right)^{\mathrm{T}} S_{z2} v_{a2} \\ &= \left(Q_{b1}^{\mathrm{T}} e_z\right)^{\mathrm{T}} v_{b1} + \left(Q_{b2}^{\mathrm{T}} e_z\right)^{\mathrm{T}} v_{b2}, \end{aligned} \tag{60}$$

and $v_b = S_z \dot{Q}_a$ is horizontal. Hence, if $v_a = \dot{Q}_a$ is the horizontal lift at Q_a, then $v_b = S_z \dot{Q}_a$ is the horizontal lift at Q_b. Since the standard metric (21) is bi-invariant, we have $\langle u, v \rangle = \langle S_z u, S_z v \rangle$ for any $S_z \in H_z$. The claim follows. ∎

5.4. Geodesics and the exponential map. The goal of this section is to show that the canonical projection of geodesics in $SO(3) \times SO(3)$ are geodesics in $\mathcal{M}_{\mathcal{I}}$, thus giving a simple expression for the exponential map. Proposition 2.1 tells us that to find geodesics in $\mathcal{M}_{\mathcal{I}}$, we can focus on finding geodesics in $SO(3) \times SO(3)$ for which the tangent vector is always horizontal. The same idea is repeatedly used in [8] to give expressions for the geodesics in the Stiefel and Grassmann manifolds. Here we now give a direct proof that if a geodesic $Q(t) \in SO(3) \times SO(3)$ has a horizontal initial tangent vector $\dot{Q}(0)$, then the tangent is horizontal for every t.

Proposition 5.4. *Let $V \in T\big(SO(3) \times SO(3)\big)$ be a vector field of the form*

$$W(t)^\vee = \operatorname{stack}\left(R_1(t)^{\mathrm{T}} e_z, R_2(t)^{\mathrm{T}} e_z\right) \tag{61}$$

defined along a geodesic $Q(t) \in SO(3) \times SO(3)$. Then we have

$$\langle \dot{Q}(t), W(t) \rangle = \langle \dot{Q}(0), W(0) \rangle. \tag{62}$$

Proof. Denote the tangent to the geodesic $Q(t)$ as

$$\dot{Q}(t)^\vee = \operatorname{stack}(v_1, v_2) \tag{63}$$

and let

$$m(t) = \langle \dot{Q}(t), W(t) \rangle = v_1^{\mathrm{T}} R_1^{\mathrm{T}} e_z + v_2^{\mathrm{T}} R_2^{\mathrm{T}} e_z. \tag{64}$$

Taking the derivative, we have

$$\dot{m}(t) = v_1^{\mathrm{T}} [v_1]_\times{}^{\mathrm{T}} R_1^{\mathrm{T}} w_1 + v_2^{\mathrm{T}} [v_2]_\times{}^{\mathrm{T}} R_2^{\mathrm{T}} w_2 \equiv 0. \tag{65}$$

Since the first derivative of $m(t)$ is identically zero, $m(t)$ must be constant, which implies (62). ∎

Combining Propositions 2.1 and 5.4, we get that the exponential map in $\mathcal{M}_{\mathcal{I}}$, i.e.,

$$[Q_b] = \exp_{[Q_a]}(v_a),\ [Q_a] \in \mathcal{M}_{\mathcal{I}},\ v_a \in T_{[Q_a]}\mathcal{M}_{\mathcal{I}}, \tag{66}$$

is obtained by projecting the exponential map in $SO(3) \times SO(3)$:

$$Q_b = \exp_{Q_a}(\tilde{v}_a),\ Q_a \in \pi^{-1}_{\mathcal{M}_{\mathcal{I}}}([Q_a]), \tag{67}$$

where $\tilde{v}_a$ is the horizontal lift of v_a, i.e.,

$$[Q_b] = \pi_{\mathcal{M}_{\mathcal{I}}}\big(\exp_{Q_a}(\tilde{v}_a)\big). \tag{68}$$

5.5. The distance and the logarithm map. Let $Q_a = (Q_{a1}, Q_{a2})$ and $Q_b = (Q_{b1}, Q_{b2})$ be two points in $SO(3) \times SO(3)$. We would like to find the distance between $[Q_a]$ and $[Q_b]$ and the logarithm map $\log_{[Q_a]}[Q_b]$. In general, we cannot directly use the distance and logarithm map in $SO(3) \times SO(3)$ because the tangent of the corresponding geodesic is not horizontal. However, we can "move" Q_b to another representative of the equivalence class $[Q_b]$, so that the geodesic between Q_a and Q_b corresponds to a geodesic between $[Q_a]$ and $[Q_b]$.

Finding the logarithm as an optimization problem. The following result shows that the correct way to "move" Q_b is by solving a one-dimensional minimization problem.

Proposition 5.5. *Define the cost*

$$f(t) = f_1(t) + f_2(t), \qquad f_i = \frac{1}{2}\theta_i^2(t), \qquad \theta_i(t) = d(Q_{ai}, R_z(t) Q_{bi}), \qquad i \in 1, 2, \tag{69}$$

and let $t_{opt} = \operatorname{argmin}_t f(t)$. Then the logarithm

$$\log_{Q_a}\big(S_z(t_{opt}) Q_b\big) = \operatorname{stack}\left(\left\{\operatorname{Log}\left(Q_{ai}^{\mathrm{T}} R_z(t_{opt}) Q_{bi}\right)\right\}_{i=1,2}\right) \tag{70}$$

is a horizontal vector in $T_{HQ}\big(SO(3) \times SO(3)\big)$.

Using (2) and the isometry given by horizontal lifts, the distance between the two elements in $\mathcal{M}_{\mathcal{I}}$ is then given by

$$d\big([Q_a],[Q_b]\big) = \|\log_{[Q_a]}[Q_b]\| = \|\log_{Q_a}\big(S_z(t_{\text{opt}})Q_b\big)\|. \tag{71}$$

Intuitively, this distance is the least amount of rotation needed to align two epipolar configurations after aligning their baselines.

Proof. We will need the following result:

$$\frac{\mathrm{d}}{\mathrm{d}t}R_z(t)Q_{bi} = R_z(t)[e_z]_\times Q_{bi} = R_z(t)Q_{bi}[Q_{bi}^{\mathrm{T}}e_z]_\times,\ i=1,2, \tag{72}$$

which in local coordinates becomes

$$\left(\frac{\mathrm{d}}{\mathrm{d}t}R_z(t)Q_{bi}\right)^\vee = Q_{bi}^{\mathrm{T}}e_z. \tag{73}$$

Taking the derivative of each term f_i, we have

$$\begin{aligned}\dot f_i(t) &= -\operatorname{Log}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi})^{\mathrm{T}}Q_{bi}^{\mathrm{T}}e_z = -\operatorname{Log}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi})^{\mathrm{T}}Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}Q_{bi}^{\mathrm{T}}e_z\\ &= -e_z^{\mathrm{T}}Q_{ai}\operatorname{Log}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}),\end{aligned} \tag{74}$$

where we used the fact that $R^{\mathrm{T}}\operatorname{Log}(R) = \operatorname{Log}(R)$ and, similarly, $R_z(t)e_z = e_z$. For $t = t_{\text{opt}}$, we have $\dot f_1(t_{\text{opt}}) + \dot f_2(t_{\text{opt}}) = 0$, which, together with (53), implies that the vector (70) is in the horizontal space at Q_a. ∎

Solving the optimization problem. The problem now is to find t_{opt}, the minimizer of f. In general, this is a nonlinear optimization problem with multiple local minima (see Figure 2 for an example). However, we can exploit its special structure (continuous, periodic, and piecewise convex) to reliably and efficiently find the global minimizer t_{opt}.

First, let us consider each function f_i separately. Using (2), the derivative of f_i is given by

$$\dot f_i(t) = e_z^{\mathrm{T}}Q_{ai}\operatorname{Log}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}) = \theta_i(t)e_z^{\mathrm{T}}Q_{ai}u_i, \tag{75}$$

where (using the closed-form expression of Log from [17])

$$u_i = \frac{1}{2\sin\theta_i(t)}\left[(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}) - (Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi})^{\mathrm{T}}\right]_\times^{\text{inv}} \tag{76}$$

is the normalized version of the logarithm vector. Notice that the derivative of f exists everywhere except at a point t_{di} for which $\sin\big(\theta_i(t_{di})\big) = 0$. The following proposition gives a way to compute the location of this point. We use the notation $(A)_{i,j}$ to denote the i,jth element of a matrix A.

Proposition 5.6. *Let θ_i be defined as in (69) and define*

$$c_{1i} = (Q_{bi}Q_{ai}^{\mathrm{T}})_{1,1} + (Q_{bi}Q_{ai}^{\mathrm{T}})_{2,2}, \qquad c_{2i} = (Q_{bi}Q_{ai}^{\mathrm{T}})_{1,2} - (Q_{bi}Q_{ai}^{\mathrm{T}})_{2,1} \tag{77}$$

$$\phi_i = \operatorname{arctan}_2(c_{1i}, c_{2i}). \tag{78}$$

Then the function $\theta_i(t)$ is continuous, 2π-periodic and

$$\sin\big(\theta_i(t_{di})\big) = 0 \quad \textit{for} \quad t_{di} = \frac{3}{2}\pi - \phi_i. \tag{79}$$

For the proof of this proposition, we will need the following lemma.

Lemma 5.7. *Define c_{1i}, c_{2i} as in (77) and let*

$$c_{3i} = (Q_{bi}Q_{ai}^{\mathrm{T}})_{3,3}. \tag{80}$$

Then the following identity holds:

$$c_{1i}^2 + c_{2i}^2 = (1 + c_{3i})^2. \tag{81}$$

Notice that this lemma is valid for any rotation $R = Q_{bi}Q_{ai}^{\mathrm{T}}$. It can be proved by parametrizing R with Euler angles, expanding (81), and then simplifying the resulting terms.

Proof of Proposition 5.6. Recall that

$$\theta_i = \arccos\left(\frac{\mathrm{tr}\big(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}\big) - 1}{2}\right). \tag{82}$$

Since the argument of arccos is a continuous function of $\cos(t)$ and $\sin(t)$ alone (which are 2π-periodic) and since arccos is continuous on its domain, then θ_i is continuous and 2π-periodic. Now let $m_i = \sqrt{c_{1i}^2 + c_{2i}^2}$. Using the standard trigonometric identity

$$c_{1i}\cos t + c_{2i}\sin t = m_i \sin(t + \phi_i), \tag{83}$$

Lemma 5.7 and expanding the definition of $R_z(t)$ in terms of $\cos t$ and $\sin t$, one can verify that

$$\mathrm{tr}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi}) = \mathrm{tr}(R_z(t)Q_{bi}Q_{ai}^{\mathrm{T}}) = c_{1i}\cos t + c_{2i}\sin t + c_{3i} = m_i\sin(t+\phi_i) + m_i - 1. \tag{84}$$

Note that $\sin\big(\theta(t_{di})\big) = 0$ such that $\cos\big(\theta(t_{di})\big) = -1$. From (82), then, we have

$$m_i \sin(t_{di} + \phi_i) + m_i - 2 = -2 \tag{85}$$

$$\sin(t_{di} + \phi_i) = -1 \tag{86}$$

$$t_{di} + \phi_i = \frac{3}{2}\pi. \tag{87}$$

The result follows. ∎

Using the definition of DLog and its closed-form expression from [26], the second derivative of f_i is given by

$$\ddot{f}_i(t) = e_z^{\mathrm{T}}Q_{ai}\,\mathrm{DLog}(Q_{ai}^{\mathrm{T}}R_z(t)Q_{bi})Q_{ai}^{\mathrm{T}}e_z = (e_z^{\mathrm{T}}Q_{ai}u_i)^2 + \frac{\theta}{2}\cot\left(\frac{\theta}{2}\right)\left(1 - \big(e_z^{\mathrm{T}}Q_{ai}u_i\big)^2\right). \tag{88}$$

Note that (as a simple plot can confirm)

$$0 \le \frac{\theta}{2}\cot\left(\frac{\theta}{2}\right) \le 1 \quad \text{for } \theta \in [-\pi, \pi]. \tag{89}$$

This implies that $\ddot{f} \ge 0$ and that f is convex between discontinuity points.

Algorithm 1. Global minimization of $f(t)$.

Compute the points t_{di}, $i = 1, 2$ (assume $t_{d1} < t_{d2}$).
Define intervals $S_1 = [t_{d1}, t_{d2}]$ and $S_2 = [t_{d2}, t_{d1} + 2\pi]$.
for $i \in 1, 2$ **do**
    if $\operatorname{sign}\big(\dot{f}^+\big(\min(S_i)\big)\big) \neq \operatorname{sign}\big(\dot{f}^-\big(\max(S_i)\big)\big)$ **then**
        Compute $t_{\mathrm{opt},i} = \operatorname{argmin}_{t \in S_i} f(t)$ using the projected Newton's method.
    end if
end for
Select t_{opt} as the point $t_{\mathrm{opt},i}$ for which f is minimum.

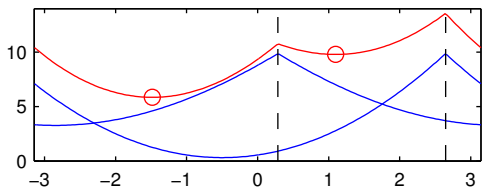


Figure 2. *An example realization of the cost $f(t)$ from (69). Blue and red lines: value of each term f_i and of f, respectively. Black dashed line: location of the discontinuity points $\{t_{di}\}$ computed using Proposition 5.6. Red circles: local minimizers $\{t_{opt,i}\}$ computed in Algorithm 1.*

In summary, from the results above, the function f is continuous, 2π-periodic and with positive second derivative except at $\{t_{di} + 2k\pi\}$, $k \in \mathbb{Z}$. Assuming (without loss of generality) the ordering $-\frac{\pi}{2} \leq t_{d1} \leq t_{d2} \leq \frac{\pi}{2}$, this suggests an algorithm to find all the global minimizers of f by considering separately the two intervals $[t_{d1}, t_{d2}]$ and $[t_{d2}, t_{d1} + 2\pi]$ (on which the function is convex and differentiable). Since we have a closed-form expression for $\ddot{f}$, we can use Newton's method (with an additional projection of the iterates to the interval). In addition, it is possible to show (using the intermediate value's theorem on $\dot{f}$) that if $\dot{f}$ has the same sign at the two extremum points of an interval, then that interval does not contain a local minimizer, and it can be skipped. These steps are summarized in Algorithm 1 (see also Figure 2). We use the notation $\dot{f}^+$ and $\dot{f}^-$ to denote right and left derivatives, respectively. Note that Algorithm 1 is only a basic version. A complete version would also consider degenerate cases, where $m_i = 0$ for some $i \in \{1, 2\}$ or where $t_{d1} = t_{d2}$. In our experiments, we saw that an interval could be skipped about 25% of the time, and that the Newton's iteration took about five to eight iterations to converge to the global minimum up to machine's precision ($2 \cdot 10^{-16}$). As a comparison, the method suggested in [24] achieves a precision of only 10^{-4} after about five iterations, and it does not guarantee global convergence (i.e., the logarithm might correspond to a nonminimal geodesic).

5.6. Curvature. As already mentioned, the curvature of the manifold plays a special role in the convergence guarantees for some optimization algorithms. For the sectional curvature of $\mathcal{M}_{\mathcal{I}}$, $K^{\mathcal{M}_{\mathcal{I}}}_{\sigma(v,w)}(x)$, we can use the fact that signed essential manifold is a submersion in $SO(3) \times SO(3)$. We can then readily obtain a simple expression by combining (27) with (12):

$$K^{\mathcal{M}_{\mathcal{I}}}_{\sigma(v,w)}(x) = \frac{1}{4}\|[u^\vee]_\times v^\vee\| + \frac{3}{4}\|\mathcal{V}[u^\vee]_\times v^\vee\|. \tag{90}$$

Note that, from the properties of projections, we have that $\|\mathcal{V}[V,W]\|^2 \leq \|[V,W]\|^2$. The curvature can then be bounded as

$$0 \leq K^{\mathcal{M}_{\mathcal{I}}}_{\sigma(v,w)}(x) \leq 2. \tag{91}$$

Hence, $\mathcal{M}_{\mathcal{I}}$ has nonnegative and bounded curvature.

6. The (unsigned) normal essential manifold. In this section, we extend the derivations performed in section, 5 to show that the original space of essential matrices $\mathcal{M}_E$ is indeed a manifold for which $\mathcal{M}_{\mathcal{I}}$ is a Riemannian covering. This will then provide us with simple expressions for the exponential and logarithm maps.

6.1. The signed normal essential manifold as a Riemannian covering. We first prove that $\mathcal{M}_E$ is indeed a manifold by seeing $\mathcal{M}_E$ as a quotient of $\mathcal{M}_{\mathcal{I}}$. This proposition is analogous to Proposition 5.2.

Proposition 6.1. *The canonical projection $\pi_{\mathcal{M}_E} : \mathcal{M}_{\mathcal{I}} \to \mathcal{M}_E$ is a smooth covering map, and $\mathcal{M}_E$ is a manifold.*

Proof. The action of S_π is smooth (since is just a matrix multiplication) and proper (again, this is automatic since $\mathcal{M}_{\mathcal{I}}$ is compact). It is also free: by way of contradiction, assume that this was not true and that there exists $S_\pi \in H_\pi$, $S_\pi \neq (I, I)$ such that $S_\pi[Q] = [Q]$, where $[Q] \in \mathcal{M}_{\mathcal{I}}$ and S_π acts on $[Q]$ component-wise. Since this is an equality between equivalence classes, for any representative $Q \in SO(3) \times SO(3)$ there must exist $S_z \in H_z$ such that $S_\pi S_z Q = Q$. By applying Q^{-1} on the right, this implies $S_\pi = S_z^{-1}$ for some $S_z \in H_z$, i.e., $S_\pi \in H_z$. But one can verify by inspection that $H_\pi \cap H_z = (I, I)$, thus contradicting $S_\pi \neq (I, I)$. Hence, the action is both proper and free. The claim then follows by applying Theorem 9.19 in [15]. ∎

Given a point $[[Q]] \in \mathcal{M}_E$, we define the induced metric on $\mathcal{M}_E$ similarly to (56), i.e.,

$$\langle \tilde{u}, \tilde{v} \rangle_{[[Q]]} = \langle u, v \rangle_{[Q]}, \tag{92}$$

where $[Q] \in \pi^{-1}_{\mathcal{M}_E} \subset \mathcal{M}_{\mathcal{I}}$ is a representative of $[[Q]]$ and $v, w \in T_{[Q]}\mathcal{M}_{\mathcal{I}}$ are $\tilde{v}, \tilde{w} \in T_{[[Q]]}\mathcal{M}_E$. We then have the following.

Proposition 6.2. *The metric $\langle , \rangle_{[[Q]]}$ defined in (92) does not depend on the choice of the representative $[Q]$.*

The proof is identical to the one of Proposition 5.3, with the only difference that now we need to consider representatives of $[Q]$ and not Q. The result, again, rests on the fact that H_π is a group of global isometries for $SO(3) \times SO(3)$.

Since H_π is discrete, given any $[[Q]] \in \mathcal{M}_E$, the vertical space of $T_{V[Q]}\mathcal{M}_{\mathcal{I}}$ at a point $[Q] \in \pi^{-1}_{\mathcal{M}_E}([[Q]])$ is simply the origin (i.e., the zero vector). Therefore, we can identify $T_{H[[Q]]}\mathcal{M}_E$ with $T_{[[Q]]}\mathcal{M}_E$ and $T_{[Q]}\mathcal{M}_{\mathcal{I}}$. In other words, every vector in $T_{[[Q]]}\mathcal{M}_E$ is horizontal and can be identified with a vector in $T_{[Q]}\mathcal{M}_{\mathcal{I}}$. From Proposition 2.1, it immediately follows

that geodesics in $\mathcal{M}_E$ are the projections of geodesics in $\mathcal{M}_{\mathcal{I}}$. The exponential map for $\mathcal{M}_E$ is then given by

$$[[Q_b]] = \exp_{[[Q_a]]}(v_a) = \pi_{\mathcal{M}_E}\big(\exp_{[Q_a]} \tilde{v}_a\big), \tag{93}$$

where $\tilde{v}_a$ is the horizontal lift (i.e., the identification) of v_a.

Since any vector in $T_{[[Q]]}\mathcal{M}_E$ is horizontal, we have four candidates for the inverse of (93), that is, the logarithm $v_a = \log_{[[Q_a]]}\big([[Q_b]]\big)$. These are given by

$$v_a^{(i)} = \log_{[Q_a]}\big(S_\pi^{(i)}[Q_b]\big), \tag{94}$$

where $S_\pi^{(i)}$, $i \in \{1, \dots, 4\}$ is an enumeration of the four elements of H_π. Since we are interested in minimal geodesics, we pick the closest candidate, that is,

$$\log_{[[Q_a]]}\big([[Q_b]]\big) = v_a^{(i_{\text{opt}})}, \qquad i_{\text{opt}} = \operatorname*{argmin}_{i\in\{1,\dots,4\}} \|v_a^{(i)}\|. \tag{95}$$

This means that we need to solve the problem described in subsection 5.5 four times, one for each candidate, and then pick the one which gives the lower distance.

To conclude, we can again use formula (12) for computing the curvature. However, since the vertical space is trivial, we have that $\mathcal{M}_E$ and $\mathcal{M}_{\mathcal{I}}$ have the same curvature, that is

$$K^{\mathcal{M}_E}_{\sigma(v,w)}([[Q]]) = K^{\mathcal{M}_{\mathcal{I}}}_{\sigma(\tilde{v},\tilde{w})}([Q]). \tag{96}$$

6.2. Comparison with previous formulations. Among the papers that use the relative pose between cameras to parametrize the essential space, the definition of normalized essential space used in [18] is compatible with $\mathcal{M}_E$, while the definition used in [22] (which includes the chirality constraints explicitly) is compatible with $\mathcal{M}_{\mathcal{I}}$. However, since the parametrization is based on the product $\mathbb{S}^2 \times SO(3)$, the resulting geodesics are not equivalent to those obtained in subsection 5.4. The more recent paper [7] includes an ad-hoc the definition of a product operation on the sphere $\mathbb{S}^2$; this is to provide an aligment between tangent spaces similar to the one given by left-invariant vector fields, and to simplify the definition of exponential and logarithm maps. At a very high level, we follow a similar process through the use of quotient manifolds. However, the results of [7] do not have an immediate geometrical interpretation in terms of epipolar configurations, while our results follow the intrinsic ambiguities of the problem.

For the papers using the parametrization derived from the SVD [9, 10, 14, 24, 25], the definition used is the same as $\mathcal{M}_{\mathcal{I}}$, and also the geodesics curves coincide. However, these papers do not fully consider the action of the group H_π and the chirality constraint. In particular, Proposition 4.3 shows that the claim made in [24] that an essential matrix E corresponds uniquely to a point in $\mathcal{M}_{\mathcal{I}}$ is false (this was already pointed out in [7]). Similarly, [9, 10, 25] only consider the action of H_z. The paper [14] is the closest to the present formulation. However, it does not consider as equivalent the essential matrices given by E and $-E$ (although the two cannot be distinguished without the chirality constraint). Therefore, it obtains a formulation where the quotient is only taken with respect to H_z and $H_{x\pi}$ (defined in (40)) instead of H_z and H_π.

Finally, none of the papers above give expressions for the curvature of the manifold.

7. Optimization on the essential manifold. In this section, we provide two examples of how our proposed Riemannian manifold structure can be used for optimization and for performing statistical operations.

7.1. Minimization of a function of E. We consider the problem of minimizing a function $f : \mathbb{R}^{3\times 3} \to \mathbb{R}$, which takes as input an essential matrix E. An example of this cost function is the Sampson error [13], which is defined as

$$f(E) = \sum_{p=1}^{P} \frac{(x_{1p}^{\mathrm{T}} E x_{2p})^2}{\sum_{k=1}^{2} (e_k^{\mathrm{T}} E x_{2p})^2 + (x_{1p}^{\mathrm{T}} E e_k)^2}, \tag{97}$$

where $\{(x_{1p}, x_{2p}\}_{p=1}^{P}$ are pairs of noisy image points that correspond to the same 3-D point in the scene. Equation (97) represents an approximation of the reprojection error, and it is one of the standard choices in optimal two-view Structure from Motion [13, 18]. We use this function as a concrete example, but the procedure in this section is general and could be applied to other functions as well.

In order to minimize the cost function f, we can combine the local parametrization of $\mathcal{M}_{\mathcal{I}}$ given by the exponential map (67) with the trust-region algorithms described in [1], which represent the state of the art in numerical optimization on manifolds. The only obstacle to this plan is that we need to consider the function $f_{\mathcal{M}_{\mathcal{I}}} : \mathcal{M}_{\mathcal{I}} \to \mathbb{R}$, $f_{\mathcal{M}_{\mathcal{I}}} = f \circ E$ given by (34) and compute the Riemannian gradient $\operatorname{grad} f_{\mathcal{M}_{\mathcal{I}}}$ and the Riemannian Hessian $\operatorname{Hess} f_{\mathcal{M}_{\mathcal{I}}}$ defined in (4) and (5) (the latter is necessary in order to use methods with quadratic asymptotic convergence rate). These quantities are not the same as the gradient $\operatorname{grad}_E f$ and Hessian $\operatorname{Hess}_E f$ obtained by considering f as a simple function of $\mathbb{R}^{3\times 3}$ (which can be computed as a standard application of multivariate calculus). Nonetheless, we will show in this section how grad and Hess can be computed from their Euclidean counterparts grad_E and Hess_E.

Consider a geodesic curve

$$Q(t) = \exp_{Q_0}(v) = (Q_1(t), Q_2(t)), \tag{98}$$

where Q_0 is arbitrary and v is horizontal. We denote its image under (34) as $E(t) \doteq E\big(Q(t)\big)$. From the definitions of exponential map in (22) and (24), one can verify that its tangent and acceleration are

$$\dot{Q} = (\dot{Q}_1, \dot{Q}_2) = (Q_1[v_1]_\times, Q_2[v_2]_\times) \tag{99}$$

$$\ddot{Q} = (Q_1[v_1]_\times^2, Q_2[v_2]_\times^2). \tag{100}$$

The corresponding quantities for $E(t)$ are

$$\dot{E} = [v_1]_\times^{\mathrm{T}} E + E[v_2]_\times = \dot{Q}_1^{\mathrm{T}} Q_1 E + E Q_2^{\mathrm{T}} \dot{Q}_2 \tag{101}$$

$$\ddot{E} = [v_1]_\times^{2\mathrm{T}} E + 2[v_1]_\times^{\mathrm{T}} E[v_2]_\times + E[v_2]_\times^2. \tag{102}$$

Considering the function $f_{\mathcal{M}_{\mathcal{I}}}(Q(t)) = f(E(t))$, from its first and second derivatives around $t = 0$ we obtain

$$\operatorname{tr}(\dot{E}^{\mathrm{T}} \operatorname{grad}_E f) = \langle \dot{Q}, \operatorname{grad} f_{\mathcal{M}_{\mathcal{I}}} \rangle \tag{103}$$

$$\operatorname{tr}(\ddot{E}^{\mathrm{T}} \operatorname{grad}_E f) + \operatorname{tr}(\dot{E}^{\mathrm{T}} \operatorname{Hess}_E f[\dot{E}]) = \langle \dot{Q}, \operatorname{Hess} f_{\mathcal{M}_{\mathcal{I}}}[\dot{Q}] \rangle. \tag{104}$$

The overall idea for the remainder of the section is to manipulate the left-hand sides of (103) and (104) to obtain the right-hand sides while keeping in mind that $\operatorname{grad} f_{\mathcal{M}_{\mathcal{I}}}$ needs to be a vector in $T_{HQ}\big(SO(3)\times SO(3)\big)$ and $\operatorname{Hess} f_{\mathcal{M}_{\mathcal{I}}}$ needs to be a symmetric map $T_{HQ}\big(SO(3)\times SO(3)\big)\to T_{HQ}\big(SO(3)\times SO(3)\big)$.

For brevity, let $G=\operatorname{grad}_E f$ and $H=\operatorname{Hess}_E f[\dot E]$. Then, starting from the left-hand side of (103), we can expand $\dot E$ and rearrange terms such that we make $\dot Q_1,\dot Q_2$ appear and obtain the sum of two inner products of the form (21) for $SO(3)\times SO(3)$. More explicitly, we have

$$
\begin{aligned}
\operatorname{tr}(\dot E^{\mathrm T}G) &= \operatorname{tr}(E^{\mathrm T}[v_1]_\times G)+\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}G)=\operatorname{tr}([v_1]_\times GE^{\mathrm T})+\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}G)\\
&=\operatorname{tr}([v_1]_\times{}^{\mathrm T}EG^{\mathrm T})+\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}G)\\
&=\operatorname{tr}\big([v_1]_\times{}^{\mathrm T}\operatorname{skew}(EG^{\mathrm T})\big)+\operatorname{tr}\big([v_2]_\times{}^{\mathrm T}\operatorname{skew}(E^{\mathrm T}G)\big)\\
&=\operatorname{tr}\big([v_1]_\times{}^{\mathrm T}Q_1^{\mathrm T}Q_1\operatorname{skew}(EG^{\mathrm T})\big)+\operatorname{tr}\big([v_2]_\times{}^{\mathrm T}Q_2^{\mathrm T}Q_2\operatorname{skew}(E^{\mathrm T}G)\big)\\
&=\operatorname{tr}\big(\dot Q_1^{\mathrm T}Q_1\operatorname{skew}(EG^{\mathrm T})\big)+\operatorname{tr}\big(\dot Q_2^{\mathrm T}Q_2\operatorname{skew}(E^{\mathrm T}G)\big),
\end{aligned}
\tag{105}
$$

where we used, in sequence, (101), the properties $\operatorname{tr}(AB)=\operatorname{tr}(BA)$, $\operatorname{tr}(A)=\operatorname{tr}(A^{\mathrm T})$, $\operatorname{tr}(A[v]_\times)=\operatorname{tr}(\operatorname{skew}(A)[v]_\times)$ (since $[v]_\times$ is antisymmetric), $Q_1^{\mathrm T}Q_1=Q_2^{\mathrm T}Q_2=I$, and (99). Note that, alternatively, one can use the second expression of $\dot E$ in (101) and the projection on the tangent spaces at Q_1 and Q_2 to obtain the same result.

Comparing subsection 7.1 with (21) for $SO(3)\times SO(3)$ and using (11), we obtain from (103) that

$$
\operatorname{grad} f_{\mathcal{M}_{\mathcal{I}}}=\mathcal{H}\big(Q_1\operatorname{skew}(EG^{\mathrm T}),Q_2\operatorname{skew}(E^{\mathrm T}G)\big).
\tag{106}
$$

Next, we consider the two terms on the left-hand side of (104) independently. For the first term, our goal is to rearrange terms to have $\dot Q_1$ and $\dot Q_2$ appear in a quadratic expression (two terms with $\dot Q_i$, $\dot Q_i^{\mathrm T}$ and two cross-terms with $\dot Q_1$, $\dot Q_2^{\mathrm T}$ and $\dot Q_1^{\mathrm T}$, $\dot Q_2$). Explicitly, we have

$$
\begin{aligned}
\operatorname{tr}(\ddot E^{\mathrm T}G) &= \operatorname{tr}(E^{\mathrm T}[v_1]_\times{}^2G)+2\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}[v_1]_\times G)+\operatorname{tr}([v_2]_\times{}^{2\mathrm T}E^{\mathrm T}G)\\
&=-\operatorname{tr}([v_1]_\times{}^{\mathrm T}GE^{\mathrm T}[v_1]_\times)+\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}[v_1]_\times G)\\
&\quad+\operatorname{tr}([v_1]_\times{}^{\mathrm T}E[v_2]_\times G^{\mathrm T})-\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}G[v_2]_\times)\\
&=-\operatorname{tr}\big([v_1]_\times{}^{\mathrm T}\operatorname{sym}(GE^{\mathrm T})[v_1]_\times\big)+\operatorname{tr}([v_2]_\times{}^{\mathrm T}E^{\mathrm T}[v_1]_\times G)\\
&\quad+\operatorname{tr}([v_1]_\times{}^{\mathrm T}E[v_2]_\times G^{\mathrm T})-\operatorname{tr}\big([v_2]_\times{}^{\mathrm T}\operatorname{sym}(E^{\mathrm T}G)[v_2]_\times\big)\\
&=-\operatorname{tr}\big([v_1]_\times{}^{\mathrm T}Q_1^{\mathrm T}Q_1\operatorname{sym}(GE^{\mathrm T})Q_1^{\mathrm T}Q_1[v_1]_\times\big)+\operatorname{tr}([v_2]_\times{}^{\mathrm T}Q_2^{\mathrm T}Q_2E^{\mathrm T}Q_1^{\mathrm T}Q_1[v_1]_\times G)\\
&\quad+\operatorname{tr}([v_1]_\times{}^{\mathrm T}Q_1^{\mathrm T}Q_1EQ_2^{\mathrm T}Q_2[v_2]_\times G^{\mathrm T})-\operatorname{tr}\big([v_2]_\times{}^{\mathrm T}Q_2^{\mathrm T}Q_2\operatorname{sym}(E^{\mathrm T}G)Q_2^{\mathrm T}Q_2[v_2]_\times\big)\\
&=-\operatorname{tr}\big(\dot Q_1^{\mathrm T}Q_1\operatorname{sym}(GE^{\mathrm T})Q_1^{\mathrm T}\dot Q_1\big)+\operatorname{tr}(\dot Q_2^{\mathrm T}Q_2E^{\mathrm T}Q_1^{\mathrm T}\dot Q_1G)\\
&\quad+\operatorname{tr}(\dot Q_1^{\mathrm T}Q_1EQ_2^{\mathrm T}\dot Q_2G^{\mathrm T})-\operatorname{tr}\big(\dot Q_2^{\mathrm T}Q_2\operatorname{sym}(E^{\mathrm T}G)Q_2^{\mathrm T}\dot Q_2\big),
\end{aligned}
\tag{107}
$$

where we used the same property of the trace operator as above, the property $[v]_\times=-[v]_\times{}^{\mathrm T}$, and the property $\operatorname{tr}(A^{\mathrm T}BA)=\operatorname{tr}(A^{\mathrm T}\operatorname{sym}(B)A)$. For the second term in the left-hand side of

(104), we can use the same computations as in subsection 7.1 with H instead of G. Putting both terms together, comparing with (21) for $SO(3) \times SO(3)$, and using (11), we obtain from (104) that

$$\begin{aligned}\operatorname{Hess} f_{\mathcal{M}_{\mathcal{I}}}[\dot{Q}] = \mathcal{H}\Big(Q_1\Big(-\operatorname{sym}(GE^{\mathrm{T}})Q_1^{\mathrm{T}}\dot{Q}_1 + EQ_2^{\mathrm{T}}\dot{Q}_2G^{\mathrm{T}} + \operatorname{skew}(EH^{\mathrm{T}})\Big),\\ Q_2\Big(-\operatorname{sym}(E^{\mathrm{T}}G)Q_2^{\mathrm{T}}\dot{Q}_2 + E^{\mathrm{T}}Q_1^{\mathrm{T}}\dot{Q}_1G + \operatorname{skew}(E^{\mathrm{T}}H)\Big)\Big).\end{aligned} \tag{108}$$

One can easily check from subsection 7.1 and from the fact that $\operatorname{Hess}_E f$ is self-adjoint that $\operatorname{Hess} f_{\mathcal{M}_{\mathcal{I}}}$ is self-adjoint too.

We have developed Matlab implementations for the computation of the Riemannian gradient (106) and Hessian (108) from their Euclidean counterparts for arbitrary cost functions. These routines have been integrated in MANOPT 2.0 [5], a toolbox implementing efficient trust-region methods on manifolds [1]. Results using this implementation to optimize the Sampson error (97) are given in subsection 7.2.

7.2. The Weiszfeld algorithm and pose averaging. In this section, we show a proof-of-concept application of the distance obtained in subsection 5.5 to the two-view Structure from Motion problem. Rather than achieving state-of-the-art reconstruction, the goal of this section is to show that distance between epipolar configurations obtained with our approach is meaningful.

In a standard pipeline, the relative pose (R, T) between two calibrated views is computed using RANSAC (see [13]):

- Extract pairs of matching image points $\{x_1^i, x_2^j\} \in \mathbb{R}^2$.
- For $i \in \{1, \dots, N\}$, select a random subset S_i of point pairs $\{x_1^i, x_2^j\}_{j \in S_i}$, estimate the essential matrix E_i, and compute its support (i.e., the number of points that approximatively satisfy the epipolar constraint).
- Compute the pose (R, T) from the matrix E_i with the largest support.

In [11] and [3], an alternative approach is suggested where instead of using RANSAC, each sample E_i is decomposed into a pose estimate (R_i, T_i), and then all the rotations $\{R_i\}$ are averaged. Toward this, they propose to minimize the cost

$$\varphi(R) = \sum_i d(R, R_i)^p, \tag{109}$$

where p varies from $p = 1$ (L1 averaging) or $p = 2$ (L2 averaging) by using the Weiszfeld algorithm, which we report in Algorithm 2 for points lying in a general Riemannian manifold $\mathcal{M}$. The authors of [3] show that the algorithm is provably convergent when the curvature of the manifold is nonnegative. This is the reason why the results of subsection 5.6 (for $\mathcal{M}_{\mathcal{I}}$) and subsection 6.1 (for $\mathcal{M}_E$) are important.

Strictly speaking, the traditional Weiszfeld algorithm refers only to the version $p = 1$, but it can be defined for any $p \geq 1$ [3]. The set I in (111) is used to take into account the fact that w_i becomes ill defined when $p = 1$ and the iterate x falls on one of the input points. Intuitively, each iteration of the algorithm maps the input points to the tangent space of the current iterate $x(t)$, takes the average (with weights given by the relative distances), and uses the resulting vector to obtain the next iterate $x(t + 1)$.

Algorithm 2. The Weiszfeld algorithm.

Require: Points $x_i \in \mathcal{M}$, $i \in \{1, \dots, N\}$.

1: Initialize $x(0)$

2: **for** $t \in \{0, \dots, N_t\}$ **do**

3: Update x using:

$$w_i(t) = d\big(x(t), x_i\big)^{p-2} \tag{110}$$

$$I(t) = \big\{i \in \{1, \dots, N\} : x(t) \neq x_i\big\} \tag{111}$$

$$x(t+1) = \exp_x\left(\frac{\sum_{i \in I} w_i(t) \log_x(x_i)}{\sum_{i \in I} w_i(t)}\right) \tag{112}$$

4: **end for**

In this section, we follow the same approach proposed by [11] and [3], but we average essential matrices instead of rotations. In practice, the only difference is the use of the definition of exp, log, and Riemannian distance for $\mathcal{M}_{\mathcal{I}}$ in Algorithm 2. Note that the approach proposed here has the immediate advantage of naturally considering both rotation and translation components together, while the approach of [11] and [3] considers only rotations. We compare the two approaches against standard RANSAC on the `fountain-P11` data set from [23], which includes the ground-truth pose for the cameras.

We used SIFT features extraction and matching [29] to find corresponding points between every possible pair of cameras. We excluded image pairs with fewer than 30 good matches (as determined using the essential matrix from the ground-truth pose). We then use the five-point algorithm [12] to generate the RANSAC samples E_i. Similarly to [11], we validate each of the solutions and keep only those that agree with three (randomly chosen) additional image points (the threshold used for the decision is the same as the one for RANSAC). We compare two versions of the Weiszfeld algorithm corresponding to the choice $\mathcal{M} = \mathcal{M}_{\mathcal{I}}$ and $\mathcal{M} = SO(3)$ by using between 1 and 50 RANSAC samples. We use $p = 1$, as the case with $p = 2$ was already shown to give inferior results in [28]. To initialize the algorithm, we evaluate the cost at every input sample and use the halfway point between the two samples with lower costs. Also, we set the number of iterations N_t to 30 (although, during preliminary tests, the algorithms usually converged in fewer than 15 iterations). As baselines for comparisons, we use the errors of the RANSAC solution after the same number of samples and after 2000 samples. Additionally, we test our and the RANSAC-based approaches followed by the optimization of the Sampson error described in subsection 7.1. In our tests, we found that the algorithm converges in about five iterations. As a quality measure, we consider the geodesic distance between estimated and ground-truth rotations. For our approach and the RANSAC-based solutions, we also consider the angle between the estimated translation direction and the ground truth. All the results are averaged across all the image pairs and 30 independent sampling realizations.

We report the results in Figure 3. As one can see, the Weiszfeld algorithm using the proposed distance on $\mathcal{M}_{\mathcal{I}}$ outperforms the corresponding version using the distance on $SO(3)$. This is in spite of the fact that the error metric considered is actually defined on $SO(3)$, and we attribute this to the fact that the distance on $\mathcal{M}_{\mathcal{I}}$ includes translations (which also show

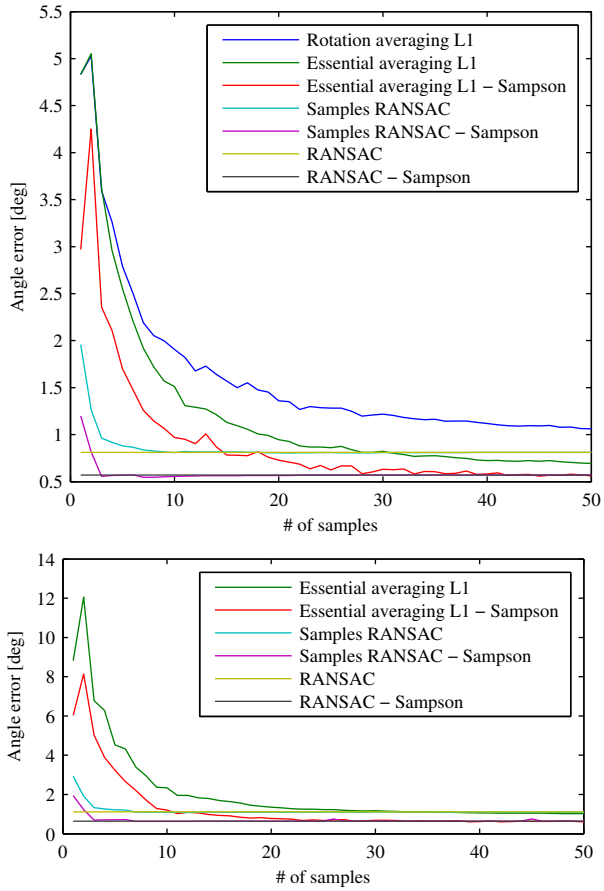


Figure 3. *Results for two-view pose estimation. Rotation (top) and translation (bottom) mean angular errors for the different methods on the* `fountain-P11` *data set.*

diminishing errors). As expected, the nonlinear refinement obtained by optimizing the Sampson error produces better results in all cases (between 0.5 and 1 degree in accuracy gain for both rotations and translations), thus validating the results of subsection 7.1. This data set also shows that, while the approach considered here gives results that are slightly better (without Sampson error optimization) or on par with the RANSAC based approach, the efficiency of the latter with a well-tuned threshold is quite hard to beat. On this data set, RANSAC reaches a good solution in fewer than 10 samples, while the Weiszfeld-based algorithm requires around 30 samples. We stress again the fact that the main purpose of this experiment is not to provide a different way to perform two-view Structure from Motion but rather to show that the distance provided by our quotient manifold representation is meaningful.

8. Conclusion. In this paper, we considered a Riemannian structure for the essential manifold and introduced a novel, geometrical interpretation that shed light on the limitations of previous approaches and on the connections with traditional concepts in computer vision. We also proposed efficient algorithms for computing the distance and logarithm map and considered an application to the problem of two-view pose estimation using averages. In our

future work, we will investigate relations between three views and determine if similar ideas can be applied to the space of trifocal tensors and other similar objects.

Appendix A. Proof of Proposition 4.2. In this section, we denote as $(A)_{i,j}$ the element in row i, column j, of a matrix A. To start the proof, we need to check that H_z and H_π are indeed groups and that their actions are well defined. The fact that H_z is a group follows from the fact that rotations around a fixed axis form a group isomorphic to $SO(2)$, and H_z is simply the Cartesian product of one of these groups with itself. The fact that H_π is a group can be checked by direct computation. Since each component of each element of H_z and H_π is a rotation, these are actually subgroups of $SO(3) \times SO(3)$, and the action defined above is well defined, mapping $SO(3)$ to itself.

We now arrive to the more involved part of the proof, showing that (39) is true. Let $S_1, S_3 \in SO(3)$ and assume that $Q \sim (S_1 Q_1, S_2 Q_2)$, i.e.,

$$E = Q_1^{\mathrm{T}} [e_z]_\times Q_2 = s Q_1^{\mathrm{T}} S_1^{\mathrm{T}} [e_z]_\times S_2 Q_2, \tag{113}$$

for some sign $s \in \{-1, 1\}$.

We will now use (113) to obtain constraints on S_1 and S_2. Using the fact that $Q_i Q_i^{\mathrm{T}} = I_3$, we have

$$[e_z]_\times = s S_1^{\mathrm{T}} [e_z]_\times S_2. \tag{114}$$

Substituting (114) into ${[e_z]_\times}^{\mathrm{T}} [e_z]_\times$ and $[e_z]_\times {[e_z]_\times}^{\mathrm{T}}$ to cancel out S_1 and S_2, respectively, and since ${[e_z]_\times}^{\mathrm{T}} [e_z]_\times = [e_z]_\times {[e_z]_\times}^{\mathrm{T}} = -{[e_z]_\times}^2 = P_z = \mathrm{diag}(1, 1, 0)$, it follows that, for both $i \in \{1, 2\}$,

$$P_z = S_i^{\mathrm{T}} P_z S_i \tag{115}$$

$$\implies S_i P_z = P_z S_i \tag{116}$$

i.e., S_i and P_z must commute. By expanding the matrix multiplications and comparing the two sides, we obtain the following constraints:

$$(S_i)_{3,1} = (S_i)_{3,2} = (S_i)_{1,3} = (S_i)_{2,3} = 0, \tag{117}$$

which is equivalent to say that S_i is of the form

$$S_i = \begin{bmatrix} * & * & 0 \\ * & * & 0 \\ 0 & 0 & * \end{bmatrix} = \mathrm{diag}(S_i', s_i). \tag{118}$$

Since $S_i \in SO(3)$, we have that $s_i \in \{1, -1\}$, $S_i' \in O(2)$, and $s_i \det(S_i) = 1$. First, consider the case $s_i = 1$. Then $\det(S_i') = 1$ and S_i can be parametrized as $S_i = R_z(\theta_i)$, i.e.,

$$S_i(\theta) = \begin{bmatrix} \cos(\theta_i) & \sin(\theta_i) & 0 \\ -\sin(\theta_i) & \cos(\theta_i) & 0 \\ 0 & 0 & 1 \end{bmatrix} = R_z(\theta_i). \tag{119}$$

Now consider the case $s_i = -1$. Then $\det(S_i') = -1$ and S_i can be parameterized as

$$S_i = \begin{bmatrix} -\cos(\theta_i) & -\sin(\theta_i) & 0 \\ -\sin(\theta_i) & \cos(\theta_i) & 0 \\ 0 & 0 & -1 \end{bmatrix} = R_y(\pi)R_z(\theta_i). \tag{120}$$

The relations (119) and (120) apply independently to S_1 and S_2, but there are additional constraints relating the two. First, notice that (114) implies

$$S_1[e_z]_\times = s[e_z]_\times S_2. \tag{121}$$

By expanding the matrix multiplication and comparing the two sides, we have eight possible cases, depending on the value of sign s and of the signs $s_1 = (S_1)_{3,3}$, $s_2 = (S_2)_{3,3}$.

1. $s = +1$, $s_1 = +1$, $s_2 = +1$: we have

$$\begin{aligned} \sin(\theta_1) &= \sin(\theta_2), & \cos(\theta_1) &= \cos(\theta_2) \\ -\cos(\theta_1) &= -\cos(\theta_2), & \sin(\theta_1) &= \sin(\theta_2) \end{aligned} \tag{122}$$

   which implies $\theta_1 = \theta_2 = \theta$. Hence,

$$S_1 = R_z(\theta), \quad S_2 = R_z(\theta). \tag{123}$$

2. $s = -1$, $s_1 = +1$, $s_2 = +1$: we have

$$\begin{aligned} \sin(\theta_1) &= -\sin(\theta_2), & \cos(\theta_1) &= -\cos(\theta_2) \\ -\cos(\theta_1) &= \cos(\theta_2), & \sin(\theta_1) &= -\sin(\theta_2) \end{aligned} \tag{124}$$

   which implies $\theta_1 = \theta_2 + \pi$. Hence,

$$S_1 = R_z(\theta), \quad S_2 = R_z(\pi)R_z(\theta). \tag{125}$$

3. $s = +1$, $s_1 = -1$, $s_2 = +1$: we have

$$\begin{aligned} -\sin(\theta_1) &= \sin(\theta_2), & \cos(\theta_1) &= \cos(\theta_2) \\ \cos(\theta_1) &= -\cos(\theta_2), & \sin(\theta_1) &= \sin(\theta_2) \end{aligned} \tag{126}$$

   which implies $\sin(\theta_i) = \cos(\theta_i) = 0$. Hence, this case is impossible.
4. $s = -1$, $s_1 = -1$, $s_2 = +1$: we have

$$\begin{aligned} -\sin(\theta_1) &= -\sin(\theta_2), & \cos(\theta_1) &= -\cos(\theta_2) \\ \cos(\theta_1) &= \cos(\theta_2), & \sin(\theta_1) &= -\sin(\theta_2) \end{aligned} \tag{127}$$

   which implies $\sin(\theta_i) = \cos(\theta_i) = 0$. Hence, this case is impossible.
5. $s = +1$, $s_1 = +1$, $s_2 = -1$: we have

$$\begin{aligned} \sin(\theta_1) &= \sin(\theta_2), & \cos(\theta_1) &= -\cos(\theta_2) \\ -\cos(\theta_1) &= -\cos(\theta_2), & \sin(\theta_1) &= -\sin(\theta_2) \end{aligned} \tag{128}$$

   which implies $\sin(\theta_i) = \cos(\theta_i) = 0$. Hence, this case is impossible.

6. $s = -1$, $s_1 = +1$, $s_2 = -1$: we have

$$\begin{aligned} \sin(\theta_1) &= -\sin(\theta_2), \quad \cos(\theta_1) = \cos(\theta_2) \\ -\cos(\theta_1) &= \cos(\theta_2), \quad \sin(\theta_1) = \sin(\theta_2) \end{aligned} \tag{129}$$

which implies $\sin(\theta_i) = \cos(\theta_i) = 0$. Hence, this case is impossible.

7. $s = +1$, $s_1 = -1$, $s_2 = -1$: we have

$$\begin{aligned} -\sin(\theta_1) &= \sin(\theta_2), \quad \cos(\theta_1) = -\cos(\theta_2) \\ \cos(\theta_1) &= -\cos(\theta_2), \quad \sin(\theta_1) = -\sin(\theta_2) \end{aligned} \tag{130}$$

which implies $\theta_1 = \theta_2 + \pi$. Hence,

$$S_1 = R_y(\pi)R_z(\theta), \quad S_2 = R_y(\pi)R_z(\pi)R_z(\theta) = R_x(\pi)R_z(\theta). \tag{131}$$

8. $s = -1$, $s_1 = -1$, $s_2 = -1$: we have

$$\begin{aligned} -\sin(\theta_1) &= -\sin(\theta_2), \quad \cos(\theta_1) = \cos(\theta_2) \\ \cos(\theta_1) &= \cos(\theta_2), \quad \sin(\theta_1) = \sin(\theta_2) \end{aligned} \tag{132}$$

which implies $\theta_1 = \theta_2 = \theta$. Hence,

$$S_1 = R_y(\pi)R_z(\theta), \quad S_2 = R_y(\pi)R_z(\theta). \tag{133}$$

Notice that four cases are impossible and that $R_z(\theta)$ appears in all the remaining cases. This represents exactly the action of $H_z H_\pi$ on $SO(3)$, and the four possible cases correspond to the four epipolar configurations in the twisted-pair ambiguity (see Figure 1).

REFERENCES

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

[2] B. Afsari, R. Tron, and R. Vidal, *On the convergence of gradient descent for finding the Riemannian center of mass*, SIAM J. Control Optim., 51 (2013), pp. 2230–2260.

[3] K. Aftab, R. Hartley, and J. Trumpf, *Generalized weiszfeld algorithms for lq optimization*, IEEE Trans. Pattern Anal. Mach. Intell., 37 (2015), pp. 728–745.

[4] M. Arie-Nachimson, S. Kovalsky, I. Kemelmacher-Shlizerman, A. Singer, and R. Basri, *Global motion estimation from point matches*, in International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission, 2012, pp. 81–88.

[5] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, *Manopt, a Matlab toolbox for optimization on manifolds*, J. Mach. Learn. Res., 15 (2014), pp. 1455–1459, http://www.manopt.org.

[6] M. P. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, MA, 1992.

[7] G. Dubbelman, L. Dorst, and H. Pijls, *Manifold statistics for essential matrices*, in IEEE European Conference on Computer Vision, 2012, pp. 531–544.

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

[9] C. Geyer and K. Daniilidis, *Mirrors in motion: Epipolar geometry and motion estimation*, in IEEE International Conference on Computer Vision, 2003, pp. 766–773.

[10] C. Geyer, S. Sastry, and R. Bajcsy, *Euclid meets Fourier: Applying harmonic analysis to essential matrix estimation in omnidirectional cameras*, in Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras, vol. 2 (2004), p. 3.

[11] R. Hartley, K. Aftab, and J. Trumpf, *L1 rotation averaging using the Weiszfeld algorithm*, in IEEE Conference on Computer Vision and Pattern Recognition, 2011.
[12] R. Hartley and H. Li, *An efficient hidden variable approach to minimal-case camera motion estimation*, IEEE Trans. Pattern Anal. Mach. Intell., (2012).
[13] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, 2004.
[14] U. Helmke, K. Hüper, P. Y. Lee, and J. Moore, *Essential matrix estimation using Gauss-Newton iterations on a manifold*, Int. J. Comput. Vis., 74 (2007), pp. 117–136.
[15] J. M. Lee, *Introduction to Smooth Manifolds*, vol. 218 of Graduate Texts in Mathematics, 2nd ed., Springer, New York, 2012.
[16] H. C. Longuet-Higgins, *A computer algorithm for reconstructing a scene from two projections*, Nature, 293 (1981), pp. 133–135.
[17] Y. Ma, *An Invitation to 3-D Vision: From Images to Geometric Models*, Springer, New York, 2004.
[18] Y. Ma, J. Košecká, and S. Sastry, *Optimization criteria and geometric algorithms for motion and structure estimation*, Int. J. Comput. Vis., 44 (2001), pp. 219–249.
[19] B. O'Neill, *The fundamental equations of a submersion.*, Michigan Math. J., 13 (1966), pp. 459–469.
[20] B. O'Neill, *Semi-Riemannian Geometry with Applications to Relativity*, 103, vol. 103, Academic Press, New York, 1983.
[21] P. Petersen, *Riemannian Geometry*, vol. 171, Springer, New York, 2006.
[22] S. Soatto, R. Frezza, and P. Perona, *Motion estimation via dynamic vision*, IEEE Trans. Automat. Control, 41 (1996), pp. 393–413.
[23] C. Strecha, W. von Hansen, L. V. Gool, P. Fua, and U. Thoennessen, *On benchmarking camera calibration and multi-view stereo for high resolution imagery*, in IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.
[24] R. Subbarao, Y. Genc, and P. Meer, *Robust unambiguous parametrization of the essential manifold*, in IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1–8.
[25] R. Subbarao and P. Meer, *Nonlinear mean shift over Riemannian manifolds*, Int. J. Comput. Vis., 84 (2009), pp. 1–20.
[26] R. Tron, *Distributed optimization on manifolds for consensus algorithms and camera network localization*, Ph.D. thesis, Johns Hopkins University, Baltimore, MD, 2012.
[27] R. Tron, B. Afsari, and R. Vidal, *Riemannian consensus for manifolds with bounded curvature*, IEEE Trans. Automat. Control, 58 (2013), pp. 921–934.
[28] R. Tron and K. Daniilidis, *On the quotient representation for the essential manifold*, in IEEE Conference on Computer Vision and Pattern Recognition, 2014.
[29] A. Vedaldi and B. Fulkerson, *VLFeat: An open and portable library of computer vision algorithms*, http://www.vlfeat.org, 2008.