SIAM J. MATRIX ANAL. APPL.
Vol. 29, No. 1, pp. 328–347

GEOMETRIC MEANS IN A NOVEL VECTOR SPACE STRUCTURE ON SYMMETRIC POSITIVE-DEFINITE MATRICES*

VINCENT ARSIGNY†, PIERRE FILLARD†, XAVIER PENNEC†, AND NICHOLAS AYACHE†

Abstract. In this work we present a new generalization of the geometric mean of positive numbers on symmetric positive-definite matrices, called Log-Euclidean. The approach is based on two novel algebraic structures on symmetric positive-definite matrices: first, a lie group structure which is compatible with the usual algebraic properties of this matrix space; second, a new scalar multiplication that smoothly extends the Lie group structure into a vector space structure. From bi-invariant metrics on the Lie group structure, we define the Log-Euclidean mean from a Riemannian point of view. This notion coincides with the usual Euclidean mean associated with the novel vector space structure. Furthermore, this means corresponds to an arithmetic mean in the domain of matrix logarithms. We detail the invariance properties of this novel geometric mean and compare it to the recently introduced affine-invariant mean. The two means have the same determinant and are equal in a number of cases, yet they are not identical in general. Indeed, the Log-Euclidean mean has a larger trace whenever they are not equal. Last but not least, the Log-Euclidean mean is much easier to compute.

Key words. geometric mean, symmetric positive-definite matrices, Lie groups, bi-invariant metrics, geodesics

AMS subject classifications. 47A64, 26E60, 53C35, 22E99, 32F45, 53C22

DOI. 10.1137/050637996

1. Introduction.

Symmetric positive-definite (SPD) matrices of real numbers appear in many contexts. In medical imaging, their use has become common during the last 10 years with the growing interest in diffusion tensor magnetic resonance imaging (DT-MRI, or simply DTI) [3]. In this imaging technique, based on nuclear magnetic resonance (NMR), the assumption is made that the random diffusion of water molecules at a given position in a biological tissue is Gaussian. As a consequence, a diffusion tensor image is an SPD matrix-valued image in which the SPD matrix associated with the current volume element (or voxel) is the covariance matrix of the local diffusion process. SPD matrices also provide a powerful framework for modeling the anatomical variability of the brain, as shown in [15]. More generally, they are widely used in image analysis, especially for segmentation, grouping, motion analysis, and texture segmentation [16]. They are also used intensively in mechanics, for example, with strain or stress tensors [4]. Last, but not least, SPD matrices are becoming a common tool in numerical analysis for generating adapted meshes to reduce the computational cost of solving partial differential equations (PDEs) in three dimensions [17].

As a consequence, there has been a growing need to carry out computations with these objects, for instance to interpolate, restore, and enhance images SPD matrices. To this end, one needs to define a complete operational framework. This

*Received by the editors August 11, 2005; accepted for publication (in revised form) by L. Reichel August 23, 2006; published electronically February 23, 2007. This work was supported by the INRIA, France.
http://www.siam.org/journals/simax/29-1/63799.html

†ASCLEPIOS Research Project, INRIA, Sophia-Antipolis, FR-06902, France (Vincent.Arsigny@Sophia.inria.fr, Pierre.Fillard@Sophia.inria.fr, Xavier.Pennec@Sophia.inria.fr, Nicholas.Ayache@Sophia.inria.fr).

is necessary to fully generalize to the SPD case the usual statistical tools or PDEs on vector-valued images. The framework of Riemannian geometry [8] is particularly adapted to this task, since many statistical tools [18] and PDEs can be generalized to this framework.

To evaluate the relevance of a given Riemannian metric, the properties of the associated notion of *mean* are of great importance. Indeed, most computations useful in practice involve averaging procedures. This is the case in particular for the interpolation, regularization, and extrapolation of SPD matrices, where mean values are implicitly computed to generate new data. For instance, the classical regularization technique based on the heat equation is equivalent to the convolution of the original data with Gaussian kernels.

Let $\mathcal{M}$ be an abstract manifold endowed with a Riemannian metric, whose associated distance is $d(.,.)$. Then the classical generalization of the Euclidean mean is given by the *Fréchet mean* (also called the *Riemannian mean*) [18, 19]. Let $(x_i)_{i=1}^N$ be N points of $\mathcal{M}$. Their Fréchet mean $\mathbb{E}(x_i)$ (possibly not uniquely defined) is defined as the point minimizing the following *metric dispersion*:

$$\mathbb{E}(x_i) = \arg\min_x \sum_{i=1}^{N} d^2(x, x_i). \tag{1.1}$$

One can directly use a Euclidean structure on square matrices to define a metric on the space of SPD matrices. This is straightforward, and in this setting, the Riemannian mean of a system of SPD matrices is their *arithmetic* mean, which is an SPD matrix since SPD matrices form a convex set. However, this mean is *not* adequate in many situations, for two main reasons. First, symmetric matrices with nonpositive eigenvalues are at a finite distance from any SPD matrix in this framework. In the case of DT-MRI, this is not physically acceptable, since this amounts to assuming that small diffusions (i.e., small eigenvalues) are much more likely than large diffusions (i.e., large eigenvalues). A priori, large and small diffusions are equally unlikely in DT-MRI, and a *symmetry with respect to matrix inversion* should be respected. In particular, a matrix and its inverse should be at the same distance from the identity. Therefore, the use of a generalization to SPD matrices of the *geometric* mean of positive numbers would be preferable, since such a mean is precisely invariant with respect to inversion.

Second, an SPD matrix corresponds typically to a *covariance matrix*. The value of its determinant is a direct measure of the dispersion of the associated multivariate Gaussian. The reason is that the volumes of associated trust regions are proportional to the square root of this determinant. But the Euclidean averaging of SPD matrices often leads to a *swelling effect*: the determinant of the Euclidean mean can be strictly larger than the original determinants. The reason is that the induced interpolation of determinants is polynomial and *not* monotonic in general. In DTI, diffusion tensors are assumed to be covariance matrices of the local Brownian motion of water molecules. Introducing more dispersion in computations amounts to introducing more diffusion, which is physically unacceptable. For illustrations of this effect, see [20, 21]. As a consequence, the determinant of a mean of SPD matrices should remain bounded by the values of the determinants of the averaged matrices.

To fully circumvent these difficulties, other metrics have been recently proposed for SPD matrices. With the affine-invariant metrics proposed in [12, 22, 23, 19], symmetric matrices with negative and null eigenvalues are at an infinite distance from any SPD matrix. The swelling effect has disappeared, and the symmetry with respect to inversion is respected. These new metrics provide an affine-invariant generalization of the geometric mean of positive numbers on SPD matrices. But the price paid for this success is a high computational burden in practice, essentially due to the curvature induced on the space of SPD matrices. This leads in many cases to slow and hard-to-implement algorithms (especially for PDEs) [12].

We propose here a new Riemannian framework on SPD matrices, which gives rise to a novel generalization of the geometric mean to SPD matrices. It fully overcomes the computational limitations of the affine-invariant framework, while conserving excellent theoretical properties. This is obtained with a new family of metrics named *Log-Euclidean*. Such metrics are particularly simple to use. They result in classical Euclidean computations in the domain of matrix logarithms. As a consequence, there is a closed form for the Log-Euclidean mean, contrary to the affine-invariant case. It results in a drastic reduction in computation time: the Log-Euclidean mean can be computed approximately 20 times faster.

The remainder of this article is organized as follows. In section 2, we recall a number of elementary properties of the space of SPD matrices. Then we proceed in section 3 to the theory of Log-Euclidean metrics which is based on two novel algebraic structures on SPD matrices: a Lie group structure and a new scalar multiplication which complements the new multiplication to obtain a new vector space structure. The definition of the Log-Euclidean mean is deduced from these new structures. Contrary to the affine-invariant mean, there is a closed form for the Log-Euclidean mean and it is simple to compute. In section 4 we highlight the resemblances and differences between affine-invariant and Log-Euclidean means. They are quite similar, since they have the same determinant, which is the classical geometric mean of the determinants of the averaged SPD matrices. They even coincide in a number of cases, and yet are different in general. We prove that Log-Euclidean means are strictly more anisotropic when averaged SPD matrices are isotropic enough.

2. Preliminaries.

We begin with a description of the fundamental properties and tools used in this work. First, we recall the elementary properties of the matrix exponential. Then we examine the general properties of SPD matrices. These properties are of two types: algebraic and differential. On the one hand, SPD matrices have algebraic properties because they are a special kind of invertible matrices, and on the other hand they can be considered globally as a smooth manifold and therefore have differential geometry properties. These properties are not independent: on the contrary, they are compatible in a profound way. This compatibility is the core of the approach developed here.

2.1. Notation.

We will use the following definitions and notation:

- $Sym_\star^+(n)$ is the space of SPD real $n \times n$ matrices.
- $Sym(n)$ is the vector space of real $n \times n$ symmetric matrices.
- $GL(n)$ is the group of real invertible $n \times n$ matrices.
- $M(n)$ is the space of real $n \times n$ square matrices.
- $\mathrm{Diag}(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix constructed with the real values $(\lambda_i)_{i \in 1 \ldots n}$ in its diagonal.
- For any square matrix M, $Sp(M)$ is the *spectrum* of M, i.e., the set of its eigenvalues.
- $\phi : E \to F$ is differentiable mapping between two smooth manifolds. Its differential at a point $M \in E$ acting on a infinitesimal displacement dM in the tangent space to E at M is written as $D_M\phi.dM$.

2.2. Matrix exponential. The exponential plays a central role in *Lie groups* (see [11, 5, 8]). We will consider here only the matrix version of the exponential, which is a tool that we extensively use in the next sections. We recall its definition and give its elementary properties. Last but not least, we give the Baker–Campbell–Hausdorff formula. It is a powerful tool that provides fine information on the structure of Lie groups around the identity. We will see in section 4 how it can be used to compare Log-Euclidean means to affine-invariant means in terms of anisotropy.

DEFINITION 2.1. *The exponential* $\exp(M)$ *of a matrix* M *is given by* $\exp(M) = \sum_{n=0}^{\infty} \frac{M^k}{k!}$. *Let* $G \in GL(n)$. *If there exists* $M \in M(n)$ *such that* $G = \exp(M)$, *then* M *is said to be a logarithm of* N.

In general, the logarithm of a real invertible matrix may not exist, and if it exists it may not be unique. The lack of existence is a general phenomenon in connected Lie groups. One generally needs *two* exponentials to reach every element [10]. The lack of uniqueness is essentially due to the influence of rotations: rotating of an angle α is the same as rotating of an angle $\alpha + 2k\pi$, where k is an integer. Since the logarithm of a rotation matrix directly depends on its rotation angles (one angle suffices in three dimensions, but several angles are necessary when $n > 3$), it is not unique. However, when a real invertible matrix has no (complex) eigenvalue on the (closed) negative real line, then it has a unique real logarithm whose (complex) eigenvalues have an imaginary part in $]-\pi, \pi[$ [2]. This particular logarithm is called *principal*. We will write $\log(M)$ for the principal logarithm of a matrix M whenever it is defined.

THEOREM 2.2. $\exp : M(n) \to GL(n)$ *is a* $\mathcal{C}^\infty$ *mapping. Its differential map at a point* $M \in M(n)$ *acting on an infinitesimal displacement* $dM \in M(n)$ *is given by*

$$D_M \exp .dM = \sum_{k=1}^{\infty} \frac{1}{k!} \left(\sum_{l=0}^{k-1} M^{k-l-1}.dM.M^l \right). \tag{2.1}$$

Proof. The smoothness of exp is simply a consequence of the uniform absolute convergence of its series expansion in any compact set of $M(n)$. The differential is obtained classically by a term by term derivation of the series defining the exponential. □

We see here that the noncommutativity of the matrix multiplication seriously complicates the differentiation of the exponential, which is much simpler in the scalar case. However, taking the trace in (2.1) yields the following.

COROLLARY 2.3. *We have the following simplification in terms of traces:*

$$\mathrm{Trace}(D_M \exp .dM) = \mathrm{Trace}(\exp(M).dM). \tag{2.2}$$

In the following we will also use this property on determinants.

PROPOSITION 2.4. *Let* $M \in M(n)$. *Then* $\det(\exp(M)) = \exp(\mathrm{Trace}(M))$.

Proof. This is easily seen in terms of eigenvalues of M. The Jordan decomposition of M [1] ensures that $\mathrm{Trace}(M)$ is the sum of its eigenvalues. But the exponential of a triangular matrix transforms the diagonal values of this matrix into their scalar exponential. The determinant of $\exp(M)$ is simply the product of its eigenvalues, which is precisely the exponential of the trace of M. □

THEOREM 2.5 (Baker–Campbell–Hausdorff formula [9] (matrix case)). *Let* $M, N \in M(n)$ *and* $t \in \mathbb{R}$. *When* t *is small enough, we have the following devel-*

opment, in which the logarithm used is the principal logarithm:

$$\begin{aligned} \log(\exp(t.M).\exp(t.N)) \;=\;& t.(M+N) + t^2/2([M,N]) \\ & + t^3/12([M,[M,N]] + [N,[N,M]]) \\ & + t^4/24([[M,[M,N]],N]) + O(t^5). \end{aligned} \tag{2.3}$$

We recall that $[M,N] = MN - NM$ *is the Lie bracket of* M *and* N.

The Baker–Campbell–Hausdorff formula shows how much $\exp(\log(M).\log(N))$ deviates from $M+N$ due to the noncommutativity of the matrix product. Remarkably, this deviation can be expressed only in terms of Lie brackets between M and N [14].

2.3. Algebraic properties. SPD matrices have remarkable algebraic properties. First, there always exists a unique real and *symmetric* logarithm for any SPD matrix, which is its principal logarithm. Second, if the space of SPD matrices is not a subgroup of $GL(n)$, it is stable with respect to inversion. Moreover, its spectral decomposition is particularly simple.[1]

THEOREM 2.6. *For any* $S \in Sym(n)$, *there exists an orthonormal coordinate system in which* S *is diagonal. This is particularly the case for SPD matrices.* $Sym_\star^+(n)$ *is not a subgroup of* $GL(n)$, *but it is stable by inversion. Moreover, the matrix exponential* $\exp : Sym(n) \rightarrow Sym_\star^+(n)$ *is one-to-one.*

Proof. For a proof of the first assertion, see elementary linear algebra manuals, or [1]. For the second assertion, we see from section 2.2 that SPD matrices have a unique real logarithm whose eigenvalues have an imaginary part between $-\pi$ and $+\pi$, since the eigenvalues of SPD matrices are real and always positive. The principal logarithm of an SPD matrix can be obtained simply by replacing its eigenvalues with their natural logarithms, which shows that this logarithm is symmetric. $\square$

Thanks to the existence of an orthonormal basis in which an SPD matrix (resp., a symmetric matrix) is diagonal, the logarithm (resp., the exponential) has a particularly simple expression. In such a basis, taking the log (resp., the exp) is simply done by applying its scalar version to eigenvalues:

$$\begin{cases} \log(R.\mathrm{Diag}(\lambda_1,\ldots,\lambda_N).R^T) = R.\mathrm{Diag}(\log(\lambda_1),\ldots,\log(\lambda_N)).R^T, \\ \exp(R.\mathrm{Diag}(\lambda_1,\ldots,\lambda_N).R^T) = R.\mathrm{Diag}(\exp(\lambda_1),\ldots,\exp(\lambda_N)).R^T. \end{cases}$$

These formulae provide a particularly efficient method to calculate the logarithms and exponentials of symmetric matrices, whenever the cost of a diagonalization is less than that of the many matrix multiplications (in the case of the exponential) and inversions (in the case of the logarithm) used in the general matrix case by classical algorithms [13, 24]. For small values of n, and in particular $n = 3$, we found such formulae to be extremely useful.

2.4. Differential properties. From the point of view of topology and differential geometry, the space of SPD matrices also has many particularities. The properties recalled here are elementary and will not be detailed. See [25] for complete proofs.

PROPOSITION 2.7. $Sym_\star^+(n)$ *is an open convex half-cone of* $Sym(n)$ *and is therefore a submanifold of* $Sym(n)$, *whose dimension is* $n(n+1)/2$.

2.5. Compatibility between algebraic and differential properties. We have seen that exp is a smooth bijection. We show here that the logarithm, i.e.,

[1] This is due to the fact that SPD matrices are *normal operators*, like rotations and antisymmetric matrices [1].

its inverse, is also smooth. As a consequence, all the algebraic operations on SPD matrices presented before are also smooth, in particular the inversion. Thus, the two structures are fully compatible.

THEOREM 2.8. $\log : Sym_\star^+(n) \to Sym(n)$ *is* $\mathcal{C}^\infty$. *Thus,* exp *and its inverse* log *are both smooth, i.e., they are diffeomorphisms. This is due to the fact that the differential of* exp *is nowhere singular.*

Proof. In fact, we need only prove the last assertion. If it is true, the implicit function theorem [6] applies and ensures that log is also smooth. Since the differential of exp at 0 is simply given by the identity, it is invertible by continuity in a neighborhood of 0. We now show that this propagates to the entire space $Sym(n)$. Indeed, let us then suppose that for a point M, the differential $D_{M/2}\exp$ is invertible. We claim that then $D_M \exp$ is also invertible, which suffices to prove the point. To show this, let us take $dM \in Sym(n)$ such that $D_M \exp .dM = 0$. If $D_M \exp$ is invertible, we should have $dM = 0$. To see this, remark that $\exp(M) = \exp(M/2).\exp(M/2)$. By differentiation and applying to dM, we get

$$D_M \exp .dM = 1/2((D_{M/2}\exp .dM).\exp(M/2) + \exp(M/2).(D_{M/2}\exp .dM)) = 0.$$

This implies by multiplication by $\exp(-M/2)$:

$$\exp(-M/2)(D_{M/2}\exp .dM).\exp(M/2) + (D_{M/2}\exp .dM) = 0.$$

Since

$$A^{-1}.\exp(B).A = \exp(A^{-1}.B.A)$$

we have also by differentiation

$$A^{-1}.D_B \exp(dB).A = D_B \exp(A^{-1}.dB.A).$$

Using this simplification and the hypothesis that $D_{M/2}\exp$ is invertible, we obtain

$$\exp(-M/2).dM.\exp(M/2) + dM = 0.$$

Let us rewrite this equation in an orthonormal basis in which M is diagonal with a rotation matrix R. Let (λ_i) be the eigenvalues of M and let $dN := R.dM.R^T$. Then we have

$$dN = -\mathrm{Diag}(\exp(-\lambda_1/2), \dots, \exp(-\lambda_N/2)).dN.\mathrm{Diag}(\exp(\lambda_1/2), \dots, \exp(\lambda_N/2)).$$

Coordinate by coordinate, this is written as:

$$\forall i, j : dN_{i,j}(1 + \exp(-\lambda_i/2 + \lambda_j/2)) = 0.$$

Hence for all $i, j :\ dN_{i,j} = 0$ which is equivalent to $dM = 0$. We are done. $\square$

COROLLARY 2.9. *In the space of SPD matrices, for all* $\alpha \in \mathbb{R}$*, the power mapping:* $S \mapsto S^\alpha$ *is smooth. In particular, this is true for the inversion mapping (i.e., when* $\alpha = -1$*).*

Proof. We have $S^\alpha = \exp(\alpha \log(S))$. The composition of smooth mappings is smooth. $\square$

3. Log-Euclidean means. In this section we focus on the construction of Log-Euclidean means. They are derived from two new structures on SPD matrices.

The first is a Lie group structure [11], i.e., an algebraic group structure that is *compatible* with the differential structure of the Space of SPD matrices. The second structure is a *vector space structure*. Indeed, one can define a *logarithmic scalar multiplication* that complements the Lie group structure to form a vector space structure on the space of SPD matrices. In this context, Log-Euclidean metrics are defined as bi-invariant metrics on the Lie group of SPD matrices. The Log-Euclidean mean is the Fréchet mean associated with these metrics. It is particularly simple to compute.

3.1. Multiplication of SPD matrices. It is not a priori obvious how one could define a multiplication on the space of SPD matrices compatible with classical algebraic and differential properties. How can one combine smoothly two SPD matrices to make a third one, in such a way that Id is still the identity and the usual inverse remains its inverse? Moreover, if we obtain a new Lie group structure, we would also like the matrix exponential to be the exponential associated with the Lie group structure which, a priori, can be different.

The first idea that comes to mind is to directly use matrix multiplication. But then the noncommutativity of matrix multiplication between SPD matrices stops the attempt: if $S_1, S_2 \in Sym_\star^+(n)$, $S_1.S_2$ is an SPD matrix (or equivalently, is symmetric) if and only if S_1 and S_2 commute. To overcome the possible asymmetry of the matrix product of two SPD matrices, one can simply take the *symmetric part* (i.e., the closest symmetric matrix in the sense of the Frobenius norm [7]) of the product and define the new product $\diamond$:

$$S_1 \diamond S_2 := \frac{1}{2}(S_1.S_2 + S_2.S_1).$$

This multiplication is smooth and conserves the identity and the inverse. But $S_1 \diamond S_2$ is not necessarily positive! Also, since the set of SPD matrices is not closed, one cannot define in general a closest SPD matrix, but only a closest symmetric *semidefinite* matrix [7].

In [12], affine-invariant distances between two SPD matrices S_1, S_2 are of the form

$$d(S_1, S_2) = \| \log(S_1^{-1/2}.S_2.S_1^{-1/2}) \|, \tag{3.1}$$

where $\|.\|$ is a Euclidean norm defined on $Sym(n)$. Let us define the following multiplication $\circledcirc$:

$$S_1 \circledcirc S_2 := S_1^{1/2}.S_2.S_1^{1/2}.$$

With this multiplication, the affine-invariant metric constructed in [12] can be interpreted then as a left-invariant metric. Moreover, this multiplication is smooth and compatible with matrix inversion and matrix exponential, and the product truly defines an SPD matrix. Everything works fine, except that it is not *associative*. This makes everything fail, because associativity is an essential requirement of group structure. Without it, many fundamental properties disappear. For Lie groups, the notion of adjoint representation no longer exists without associativity.

Theorem 2.8 points to an important fact: $Sym_\star^+(n)$ is diffeomorphic to its tangent space at the identity, $Sym(n)$. But $Sym(n)$ has an additive group structure, and to obtain a group structure on the space of SPD matrices, one can simply transport the additive structure of $Sym(n)$ to $Sym_\star^+(n)$ with the exponential. More precisely, we have the following.

DEFINITION 3.1. *Let $S_1, S_2 \in Sym_\star^+(n)$. We define their logarithmic product $S_1 \odot S_2$ by*

$$S_1 \odot S_2 := \exp(\log(S_1) + \log(S_2)). \tag{3.2}$$

PROPOSITION 3.2. *$(Sym_\star^+(n), \odot)$ is a group. The neutral element is the usual identity matrix, and the group inverse of an SPD matrix is its inverse in the matrix sense. Moreover, whenever two SPD matrices commute in the matrix sense, the logarithmic multiplication is equal to their matrix product. Last but not least, the multiplication is commutative.*

Proof. The multiplication is defined by addition on logarithms. It is therefore associative and commutative. Since $\log(\mathrm{Id}) = 0$, the neutral element is Id, and since $\log(S^{-1}) = -\log(S)$, the new inverse is the matrix inverse. Finally, we have $\exp(\log(S_1) + \log(S_2)) = \exp(\log(S_1)).\exp(\log(S_2)) = S_1.S_2$ when $[S_1, S_2] = 0$. ∎

THEOREM 3.3. *The logarithmic multiplication $\odot$ on $Sym_\star^+(n)$ is compatible with its structure of smooth manifold: $(S_1, S_2) \mapsto S_1 \odot S_2^{-1}$ is $\mathcal{C}^\infty$. Therefore, $Sym_\star^+(n)$ is given a commutative Lie group structure by $\odot$.*

Proof. $(S_1, S_2) \mapsto S_1 \odot S_2^{-1} = \exp(\log(S_1) - \log(S_2))$. But since exp and log and the addition are smooth, their composition is also smooth. By definition (see [8, page 29]), $Sym_\star^+(n)$ is a Lie group. ∎

PROPOSITION 3.4. $\exp : (Sym(n), +) \to (Sym_\star^+(n), \odot)$ *is a Lie group isomorphism. In particular, one-parameter subgroups of $Sym_\star^+(n)$ are obtained by taking the matrix exponential of those of $Sym(n)$, which are simply of the form $(t.V)_{t\in\mathbb{R}}$, where $V \in Sym(n)$. As a consequence, the Lie group exponential in $Sym_\star^+(n)$ is given by the classical matrix exponential on the Lie algebra $Sym(n)$.*

Proof. We have explicitly transported the group structure of $Sym(n)$ into $Sym_\star^+(n)$ so exp is a morphism. It is also a bijection, and thus an isomorphism. The smoothness of exp then ensures its compatibility with the differential structure.

Let us recall the definition of one-parameter subgroups. $(S(t))_{t\in\mathbb{R}}$ is such a subgroup if and only if we have for all t, s: $S(t + s) = S(t) \odot S(s) = S(s) \odot S(t)$. But then $\log(S(t+s) = \log(S(t) \odot S(s)) = \log(S(t)) + \log(S(s))$ by definition of $\odot$. Therefore $\log S(t)$ is also a one-parameter subgroup of $(Sym(n), +)$, which is necessarily of the form $t.V$, where $V \in Sym(n)$. V is the *infinitesimal generator* of $S(t)$. Finally, the exponential is obtained from one-parameter subgroups, which are all of the form $(\exp(t.V))_{t\in\mathbb{R}}$ (see [5, Chap. V]). ∎

Thus, we have given the space of SPD matrices a structure of Lie group that leaves unchanged the classical matrix notions of inverse and exponential. The new multiplication used, i.e., the logarithmic multiplication, generalizes the matrix multiplication when two SPD matrices do not commute in the matrix sense.

The associated Lie algebra is the space of symmetric matrices, which is diffeomorphic and isomorphic to the group itself. The associated Lie bracket is the null bracket: $[S_1, S_2] = 0$ for all $S_1, S_2 \in Sym(n)$.

The reader should note that this Lie group structure is, to our knowledge, new in the literature. For a space as commonly used as SPD matrices, this is quite surprising. The probable reason is that the Lie group of SPD matrices is *not* a multiplicative matrix group, contrary to most Lie groups.

3.2. Log-Euclidean metrics on the Lie group of SPD matrices. Now that we have given $Sym_\star^+(n)$ a Lie group structure, we turn to the task of exploring metrics compatible with this new structure. Among Riemannian metrics in Lie groups, *bi-invariant* metrics are the most convenient. We have the following definition.

DEFINITION 3.5. *A metric $\langle , \rangle$ defined on a Lie group G is said to be bi-invariant if for all $m \in G$, the left- and right-multiplication by m do not change distances between points, i.e., are isometries.*

THEOREM 3.6. *From* [5, *Chap.* V], *bi-invariant metrics have the following properties:*

1. *A bi-invariant metric is also invariant w.r.t. inversion.*
2. *It is bi-invariant if and only if for all $m \in G, Ad(m)$ is an isometry of the Lie algebra $\mathfrak{g}$, where $Ad(m)$ is the adjoint representation of m.*
3. *One-parameter subgroups of G are geodesics for the bi-invariant metric. Conversely, geodesics are simply given by left- or right-translations of one-parameter subgroups.*

COROLLARY 3.7. *Any metric $\langle , \rangle$ on $T_{Id}Sym_\star^+(n) = Sym(n)$ extended to $Sym_\star^+(n)$ by left- or right-multiplication is a bi-invariant metric.*

Proof. The commutativity of the multiplication implies that $Ad(Sym_\star^+(n)) = \{\mathrm{Id}\}$, which is trivially an isometry group. □

This result is striking. In general Lie groups, the existence of bi-invariant metrics is not guaranteed. More precisely, it is guaranteed if and only if the adjoint representation $Ad(G)$ is relatively compact, i.e., (the dimension is assumed finite) if the group of matrices given by $Ad(G)$ is *bounded* (see [5, Theorem V.5.3]). This is trivially the case when the group is commutative, as here, since $Ad(G) = \{e\}$, which is obviously bounded. Other remarkable cases where $Ad(G)$ is bounded are compact groups, such as rotations. But for noncompact noncommutative groups, there is in general *no* bi-invariant metric, as in the case of rigid transformations.

DEFINITION 3.8. *Any bi-invariant metric on the lie group of SPD matrices is also called a Log-Euclidean metric because it corresponds to a Euclidean metric in the logarithmic domain, as is shown in Corollary* 3.9.

COROLLARY 3.9. *Let $\langle , \rangle$ be a bi-invariant metric on $Sym_\star^+(n)$. Then its geodesics are simply given by the translated versions of one-parameter subgroups, namely,*

$$(\exp(V_1 + t.V_2))_{t\in\mathbb{R}}, \text{ where } V_1,\ V_2 \in Sym(n). \tag{3.3}$$

The exponential and logarithmic maps associated with the metric can be expressed in terms of matrix exponential and logarithms in the following way:

$$\begin{cases} \log_{S_1}(S_2) = D_{\log(S_1)} \exp .(\log(S_2) - \log(S_1)), \\ \exp_{S_1}(L) = \exp(\log(S_1) + D_{S_1} \log .L). \end{cases} \tag{3.4}$$

The scalar product between two tangent vectors V_1, V_2 at a point S is given by

$$\langle V_1, V_2\rangle_S = \langle D_S \log .V_1, D_S \log .V_2\rangle_{Id}. \tag{3.5}$$

From this equation, we get the distance between two SPD matrices:

$$d(S_1, S_2) = \| \log_{S_1}(S_2)\|_{S_1} = \| \log(S_2) - \log(S_1)\|_{Id}, \tag{3.6}$$

where $\|.\|$ is the norm associated with the metric.

Proof. Theorem 3.6 states that geodesics are obtained by translating one-parameter subgroups, and Proposition 3.4 gives the form of these subgroups in terms of the matrix exponential. By definition, the metric exponential $\exp_{S_1} : T_{S_1}Sym_\star^+(n) \to Sym_\star^+(n)$ is the mapping that associates with a tangent vector L the value at time 1 of the geodesic starting at time 0 from S_1 with an initial speed vector L. Differentiating the geodesic equation (3.3) at time 0 yields an initial vector speed equal to

$D_{V_1} \exp .V_2$. As a consequence, $\exp_{S_1}(L) = \exp(\log(S_1) + (D_{\log(S_1)} \exp)^{-1}.L)$. The differentiation of the equality $\log \circ \exp = \mathrm{Id}$ yields $(D_{\log(S_1)} \exp)^{-1} = D_{S_1} \log$. Hence we have the formula for $\exp_{S_1}(L)$. Solving in L the equation $\exp_{S_1}(L) = S_2$ provides the formula for $\log_{S_1}(S_2)$.

The metric at a point S is obtained by propagating by translation the scalar product on the tangent space at the identity. Let $L_S : Sym_\star^+(n) \to Sym_\star^+(n)$ be the logarithmic multiplication by S. We have $\langle V_1, V_2 \rangle_S = \langle D_S L_{S^{-1}}.V_1, D_S L_{S^{-1}}.V_2 \rangle$. But simple computations show that $D_S L_{S^{-1}} = D_S \log$. Hence we have (3.5). Finally, we combine (3.4) and (3.5) to obtain the (simple this time!) formula for the distance. ◻

COROLLARY 3.10. *Endowed with a bi-invariant metric, the space of SPD matrices is a flat Riemannian space: its sectional curvature* (*see* [8, *page* 107]) *is null everywhere.*

This is clear, since it is isometric to the $Sym(n)$ endowed with the Euclidean distance associated with the metric.

In [12], the metric defined on the space of SPD matrices is affine invariant. The action $\mathrm{act}(A)$ of an invertible matrix A on the space of SPD matrices is defined by

$$\forall S,\ \mathrm{act}(A)(S) = A.S.A^T.$$

Affine-invariance means that for all invertible matrices A, the mapping $\mathrm{act}(A) : Sym_\star^+(n) \to Sym_\star^+(n)$ is an isometry. This group action describes how an SPD matrix, assimilated to a covariance matrix, is affected by a general affine change of coordinates.

Here, the Log-Euclidean Riemannian framework will not yield full affine-invariance. However, it is not far from it, because we can obtain invariance by similarity (isometry plus scaling).

PROPOSITION 3.11. *We can endow $Sym_\star^+(n)$ with a similarity-invariant metric, for instance, by choosing $\langle V_1, V_2 \rangle := Trace(V_1.V_2)$ for $V_1, V_2 \in Sym(n)$.*

Proof. Let $R \in SO(n)$ be a rotation and $s > 0$ be a scaling factor. Let S be an SPD matrix. V is transformed by the action of $s.R$ into $\mathrm{act}(sR)(S) = s^2.R.S.R^T$. From (3.6), the distance between two SPD matrices S_1 and S_2 transformed by sR is

$$d(\mathrm{act}(sR)(S_1), \mathrm{act}(sR)(S_2)) = \mathrm{Trace}(\{\log(\mathrm{act}(sR)(S_1)) - \log(\mathrm{act}(sR)(S_2))\}^2).$$

A scaling by a positive factor λ on an SPD matrix corresponds to a translation by $\log(\lambda).\mathrm{Id}$ in the domain of logarithms. Furthermore, we have $\log(R.S.R^T) = R.\log(S).R^T$ for any SPD matrix S and any rotation R. Consequently, the scaling zeros out in the previous formula and we have

$$\begin{aligned} d(\mathrm{act}(sR)(S_1), \mathrm{act}(sR)(S_2)) &= \mathrm{Trace}(\{R.(\log(S_1) - \log(S_2)).R^T\}^2) \\ &= \mathrm{Trace}(\{\log(S_1) - \log(S_2)\}^2) \\ &= d(S_1, S_2). \end{aligned}$$

Hence we have the result. ◻

Thus, we see that the Lie group of SPD matrices with an appropriate Log-Euclidean metric has many *invariance properties*: Lie group bi-invariance and similarity-invariance. Moreover, Theorem 3.6 shows that the inversion mapping $S \mapsto S^{-1}$ is an isometry.

3.3. A vector space structure on SPD matrices. We have already seen that the Lie group of SPD matrices is isomorphic and diffeomorphic to the additive group of symmetric matrices. We have also seen that with a Log-Euclidean metric, the Lie group of SPD matrices is also isometric to the space of symmetric matrices endowed with the associated Euclidean metric. There is more: the Lie group isomorphism exp from the Lie algebra of symmetric matrices to the space of SPD matrices can be smoothly extended into an isomorphism of vector spaces. Indeed, let us define the following operation.

DEFINITION 3.12. *The logarithmic scalar multiplication* $\circledast$ *of an SPD matrix by a scalar* $\lambda \in \mathbb{R}$ *is*

$$\lambda \circledast S = \exp(\lambda . \log(S)) = S^{\lambda}. \tag{3.7}$$

When we assimilate the logarithmic multiplication to an addition and the logarithmic scalar multiplication to a usual scalar multiplication, we have all the properties of a vector space. By construction, the mapping $\exp : (Sym(N), +, .) \rightarrow (Sym_{\star}^{+}(n), \odot, \circledast)$ is a vector space isomorphism. Since all algebraic operations on this vector space are smooth, this defines what could be called a "Lie vector space structure" on SPD matrices.

Of course, this result does not imply that the space of SPD matrices is a vector subspace of the vector space of square matrices. But it shows that we can view this space as a vector space when we identify an SPD matrix with its logarithm. The question of whether or not the SPD matrix space is a vector space depends on the vector space structure we are considering, and *not* on the space itself.

From this point of view, bi-invariant metrics on the Lie group of SPD matrices are simply the classical Euclidean metrics on the vector space $(Sym(n), +, .)$. Thus, we have in fact defined a new Euclidean structure on the space of SPD matrices by transporting that of its Lie algebra $Sym(n)$ on SPD matrices. But this Euclidean structure does not have the defects mentioned in the introduction of this article: matrices with null eigenvalues are at infinite distance and the symmetry principle is respected. Last but not least, with an appropriate metric, similarity-invariance is also guaranteed.

3.4. Log-Euclidean mean. We present here the definition of the Log-Euclidean mean of SPD matrices and its invariance properties.

THEOREM 3.13. *Let* $(S_i)_1^N$ *be a finite number of SPD matrices. Then their Log-Euclidean Fréchet mean exists and is unique. It is given explicitly by*

$$\mathbb{E}_{LE}(S_1, \ldots, S_N) = \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(S_i)\right). \tag{3.8}$$

The Log-Euclidean mean is similarity-invariant, invariant by group multiplication and inversion, and is exponential-invariant (i.e., invariant with respect to scaling in the domain of logarithms).

Proof. When one expresses distances in the logarithm domain, one is faced with the classical computation of an Euclidean mean. Hence we have the formula by mapping back the results with exp in the domain of SPD matrices. Now, this mean does not depend on the chosen Log-Euclidean metric, and since there exist similarity-invariant metrics among Log-Euclidean metrics, this property propagates to the mean. The three last invariance properties are reformulations in the domain of SPD matrices of classical properties of the arithmetic mean in the domain of logarithms. $\square$

TABLE 4.1

Comparison between affine-invariant and Log-Euclidean metrics. Note on the one hand the important simplifications in terms of distance and geodesics in the Log-Euclidean case. On the other hand, this results in the use of the differentials of the matrix exponential and logarithm in the exponential and logarithm maps.

Affine-invariant metrics	Log-Euclidean metrics
Exponential map: $\exp_{S_1}(L) =$	
$S_1^{1/2}.\exp(S_1^{-1/2}.L.S_1^{-1/2}).S_1^{1/2}$	$\exp(\log(S_1) + D_{S_1}\log .L)$
Logarithm map: $\log_{S_1}(S_2) =$	
$S_1^{1/2}.\log(S_1^{-1/2}.S_2.S_1^{-1/2}).S_1^{1/2}$	$D_{\log(S_1)}\exp .(\log(S_2) - \log(S_1))$
Dot product: $\langle L_1, L_2\rangle_S =$	
$\langle S^{-1/2}.L_1.S^{-1/2}, S^{-1/2}.L_2.S^{-1/2}\rangle_{\mathrm{Id}}$	$\langle D_S\log .L_1, D_S\log .L_2\rangle_{\mathrm{Id}}$
Distance: $d(S_1, S_2) =$	
$\\| \log(S_1^{-1/2}.S_2.S_1^{-1/2})\\|$	$\\| \log(S_2) - \log(S_1)\\|$
Geodesic between S_1 and S_2:	
$S_1^{1/2}.\exp(tW).S_1^{1/2}$ with $W = \log\left(S_1^{-1/2}.L.S_1^{-1/2}\right)$	$\exp\left((1-t)\log(S_1) + t\log(S_2)\right)$
Invariance properties	
Affine-invariance	Lie group bi-invariance, Similarity-invariance

4. Comparison with the affine-invariant mean. In this section we compare the Log-Euclidean mean to the recently introduced affine-invariant mean [12, 19, 23, 22]. To this end, we first recall the differences between affine-invariant metrics and Log-Euclidean metrics in terms of elementary operators, distance, and geodesics. Then we turn to a study of the algebraic properties of Fréchet means in the Log-Euclidean and affine-invariant cases.

4.1. Elementary metric operations and invariance. Distances, geodesics, and Riemannian means take a much simpler form in the Log-Euclidean than in the affine-invariant case. Invariance properties are comparable: some Log-Euclidean metrics are not only bi-invariant but also similarity invariant. These properties are summarized in Table 4.1. However, we see in this table that the exponential and logarithmic mappings are complicated in the Log-Euclidean case by the use of the differentials of the matrix exponential and logarithm. This is the price to pay to obtain simple distances and geodesics. Interestingly, using spectral properties of symmetric matrices, one can obtain a closed form for the differential of both matrix logarithm and exponential and it is possible compute them very efficiently. See [26] for more details.

4.2. Affine-invariant means. Let $(S_i)_{i=1}^N$ be a system of SPD matrices. Contrary to the Log-Euclidean case, there is in general no closed form for the affine-invariant Fréchet mean $E_{Aff}(S_1,\dots,S_N)$ associated with affine-invariant metrics. The affine-invariant mean is defined *implicitly* by a *barycentric equation*, which is the following:

$$\sum_{i=1}^{N} \log(\mathbb{E}_{Aff}(S_1,\dots,S_N)^{-1/2}.S_i.\mathbb{E}_{Aff}(S_1,\dots,S_N)^{-1/2}) = 0. \tag{4.1}$$

This equation is equivalent to the following other barycentric equation, given in [19]:

$$\sum_{i=1}^{N} \log(\mathbb{E}_{Aff}(S_1, \ldots, S_N)^{-1}.S_i) = 0. \tag{4.2}$$

The two equations are equivalent simply because for all i,

$$\mathbb{E}_{Aff}(S_1, \ldots, S_N)^{-1/2}.S_i.\mathbb{E}_{Aff}(S_1, \ldots, S_N)^{-1/2} = A.\mathbb{E}_{Aff}(S_1, \ldots, S_N)^{-1}.S_i.A^{-1}$$

with $A = \mathbb{E}_{Aff}(S_1, \ldots, S_N)^{-1/2}$. The fact that $\log(A.S.A^{-1}) = A.\log(S).A^{-1}$ suffices to conclude.

To solve (4.1), the only known strategy is to resort to an iterative numerical procedure, such as the Gauss–Newton gradient descent method described in [12].

4.3. Geometric interpolation of determinants. The definition of the Log-Euclidean mean given by (3.8) is extremely similar to that of the classical scalar geometrical mean. We have the following classical definition.

DEFINITION 4.1. *The geometrical mean of positive numbers $d_1, \ldots, d_N$, is given by*

$$\mathbb{E}(d_1, \ldots, d_N) = \exp\left(\frac{1}{N}\sum_{i=1}^{N} \log(d_i)\right).$$

The Log-Euclidean and affine-invariant Fréchet means can *both* be considered as generalizations of the geometric mean. Indeed, their determinants are both equal to the scalar geometric mean of the determinants of the original SPD matrices. This fundamental property can be thought of as the *common property that should have all generalizations of the geometric mean* to SPD matrices.

THEOREM 4.2. *Let $(S_i)_{i=1}^N$ be N SPD matrices. Then the determinant of their Log-Euclidean and affine-invariant means is the geometric mean of their determinants.*

Proof. From Proposition 2.4 we know that $\det(\exp(M)) = \exp(\mathrm{Trace}(M))$ for any square matrix M. Then for the geometric mean, we get

$$\begin{aligned}
\det(\mathbb{E}_{LE}(S_1, \ldots, S_N)) &= \exp(\mathrm{Trace}(\log(\mathbb{E}_{LE}(S_1, \ldots, S_N)))) \\
&= \exp\left(\mathrm{Trace}\left(\frac{1}{N}\sum_{i=1}^{N}\log(S_i)\right)\right) \\
&= \exp\left(\frac{1}{N}\sum_{i=1}^{N}\log(\det(S_i))\right) \\
&= \exp\left(\mathbb{E}(\log(\det(S_1, \ldots, S_N)))\right).
\end{aligned}$$

For affine-invariant means, there is no closed form for the mean. *But* there is the barycentric equation given by (4.1). By applying the same formula as before after having taken the exponential and using $\det(S.T) = \det(S).\det(T)$ we obtain the result. ∎

Theorem 4.2 shows that the Log-Euclidean and affine-invariant means of SPD matrices are quite similar. In terms of interpolation, this result is satisfactory, since it implies that the interpolated determinant, i.e., the volume of the associated interpolated ellipsoids, will vary between the values of the determinants of the source SPD matrices. Indeed, we have the following.

Corollary 4.3. *Let $(S_i)_{i=1}^N$ be N SPD matrices. Then the determinant of their Log-Euclidean and affine-invariant means are within the interval*

$$\left[\inf_{i\in 1\ldots N}(S_i), \sup_{i\in 1\ldots N}(S_i)\right].$$

Proof. This is simply a consequence of the monotonicity of the scalar exponential and of the scalar integral. ∎

Corollary 4.4. *Let S_1 and S_2 be two SPD matrices. The geodesic interpolations provided by the affine-invariant and Log-Euclidean metrics lead to a geometric interpolation of determinants. As a consequence, this interpolation of determinants is monotonic.*

Proof. Indeed, in both cases, the interpolated determinant $\mathrm{Det}(t)$ is the geometric mean of the two determinants, i.e., at $t \in [0,1]$: $\mathrm{Det}(t) = \exp((1-t)\log(\det(S_1)) + t\log(\det(S_2)))$. This interpolation is monotonic, since the differentiation yields

$$\frac{d}{dt}\mathrm{Det}(t) = \mathrm{Det}(t)\log(\det(S_2.S_1^{-1})).$$

As a consequence, $\mathrm{Det}(t)$ is equal to $\det(S_1).\exp(t.\log(\det(S_2.S_1^{-1})))$, and the sign of $\frac{d}{dt}\mathrm{Det}(t)$ is constant and given by $\log(\det(S_2.S_1^{-1}))$. ∎

4.4. Criterion for the equality of the two means. In general, Log-Euclidean and affine-invariant means are similar, yet they are *not* identical. Nonetheless, there are a number of cases where they are identical, for example, when the logarithms of averaged SPD matrices all commute with one another. In fact, we have more as follows.

Proposition 4.5. *Let $(S_i)_{i=1}^N$ be N SPD matrices. If the Euclidean mean of the associated logarithms commutes with all $\log(S_i)$, then the Log-Euclidean and the affine-invariant means are identical.*

Proof. Let $\bar{L} := \frac{1}{N}\sum_{i=1}^N \log(S_i)$. The hypothesis is that $[\bar{L}, \log(S_i)] = 0$ for all i. This implies that $\log(\exp(-\frac{1}{2}\bar{L}).S_i.\exp(-\frac{1}{2}\bar{L})) = \log(S_i) - \bar{L}$ for all i. We see then that $\exp\bar{L}$, i.e., the Log-Euclidean mean, is the solution of (4.1), i.e., is the affine-invariant mean. ∎

So far, we have not been able to prove the converse part of this proposition. However, the next subsection provides a partial proof, valid when SPD matrices are isotropic enough, i.e., close to a scaled version of the identity. The intensive numerical experiments we have carried out strongly suggest that the result given in the next section is true in general. The full proof of this assertion will be the subject of future work.

4.5. Larger anisotropy in Log-Euclidean means. In section 4.6, we will verify experimentally that affine-invariant means tend to be less anisotropic than Log-Euclidean means. The following theorem accounts for this phenomenon when SPD matrices are isotropic enough.

Theorem 4.6. *Let $(S_i)_{i=1}^N$ be a finite number of SPD matrices close enough to the identity, so that we can apply the Baker–Campbell–Hausdorff formula in all cases (see section 2). When the logarithm of the Log-Euclidean mean does not commute with all $\log(S_i)$, then we have the following inequality:*

$$\mathit{Trace}(\mathbb{E}_{Aff}(S_1,\ldots,S_N)) < \mathit{Trace}(\mathbb{E}_{LE}(S_1,\ldots,S_N)). \tag{4.3}$$

Proof. The idea is to see how the two means differ close to the identity. To this end, we introduce a small scaling factor t and see how the two means vary when t is close to zero. For all i, let $S_{i,t}$ be the version of S_i scaled by t in the logarithmic domain. Around the identity, we can use the Baker–Campbell–Hausdorff formula to simplify the barycentric equation (4.1). Let us denote both Riemannian cases as $\mathbb{E}(S_t) = \mathbb{E}(S_{1,t}, \ldots, S_{N,t})$ and $\mathbb{E}(S) := \mathbb{E}(S_1, \ldots, S_N)$. We will also use the following notation: $\log(S_i) := L_i$, $\bar{L}_{t;Aff} := \log(\mathbb{E}_{Aff}(S_t))$ and $\bar{L}_{LE} := \log(\mathbb{E}_{LE}(S))$.

%pagebreak

First, we use twice the Baker–Campbell–Hausdorff formula to obtain the following approximation:

(4.4)
$$\begin{aligned}\log(\mathbb{E}_{Aff}(S_t)^{-1/2}.S_{i,t}.\mathbb{E}_{Aff}(S_t)^{-1/2}) = \quad & tL_i - \bar{L}_{t;Aff} - t^3 \tfrac{1}{12}[L_i, [L_i, \bar{L}_{t;Aff}]] \\ & + t^3 \tfrac{1}{24}[\bar{L}_{t;Aff}, [\bar{L}_{t;Aff}, L_i]] + O(t^5).\end{aligned}$$

Then we average over i to obtain the following approximation lemma.

LEMMA 4.7. *When t is small enough, we have:*

$$\bar{L}_{t;Aff} = t\bar{L}_{LE} + \frac{t^3}{12.N} \sum_{i=1}^{N} [L_i, [\bar{L}_{LE}, L_i]] + O(t^5). \tag{4.5}$$

Proof. To obtain the approximation, note that the second factor $t^3 \frac{1}{24}[\bar{L}_{t;Aff}, [\bar{L}_{t;Aff}, L_i]]$ in (4.4) becomes an $O(t^5)$. Indeed, when the sum over i is done, L_i becomes $\bar{L}_{LE}$. But we can replace $\bar{L}_{LE}$ with its value in term of the affine-invariance mean by using (4.4). Then, using the fact that $[\bar{L}_{t;Aff}, \bar{L}_{t;Aff}] = 0$ we see that we obtain an $O(t^5)$.

Note also that, thanks to the symmetry with respect to inversion, $\bar{L}_{t;Aff}$ becomes $-\bar{L}_{t;Aff}$ when t is changed into $-t$, i.e., $t \mapsto \bar{L}_{t;Aff}$ is odd. As a consequence, only odd terms appear in the development in powers of t. $\square$

Next, we take the exponential of (4.5) and differentiate the exponential to obtain

$$\mathbb{E}_{Aff}(S_t) = \mathbb{E}_{LE}(S_t) + D_{t\bar{L}_{LE}} \exp . \left(\frac{t^3}{12.N} \sum_{i=1}^{N} [L_i, [\bar{L}_{LE}, L_i]] \right) + O(t^5).$$

Then we use several properties to approximate the trace of affine-invariant means. First, we use Corollary 2.3 to simplify the use of the differential of the exponential. Then we approximate the exponential by the first two terms of its series expansion. We obtain

$$\mathrm{Trace}(\mathbb{E}_{Aff}(S_t)) = \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) + t^3.F(t, L_i, \bar{L}_{LE}) + O(t^5),$$

with $F(t, L_i, \bar{L}_{LE}) = \mathrm{Trace}(\exp(t\bar{L}_{LE}).\frac{1}{12.N}\sum_{i=1}^{N}[L_i, [\bar{L}_{LE}, L_i]])$. This expression can be simplified as follows:

$$\begin{aligned} F(t, L_i, \bar{L}_{LE}) &= \mathrm{Trace}\left((\mathrm{Id} + t\bar{L}_{LE}).\frac{1}{12.N} \sum_{i=1}^{N} [L_i, [\bar{L}_{LE}, L_i]] \right) + O(t^2) \\ &= \frac{t}{12.N} \sum_{i=1}^{N} \mathrm{Trace}\left(\bar{L}_{LE}.[L_i, [\bar{L}_{LE}, L_i]] \right) + O(t^2) \\ &= -\frac{t}{12.N} \sum_{i=1}^{N} \mathrm{Trace}\left(L_i^2.\bar{L}_{LE}^2 - (L_i.\bar{L}_{LE})^2 \right) + O(t^2). \end{aligned}$$

As a consequence, the difference between the two traces can be written as

$$\mathrm{Trace}(\mathbb{E}_{Aff}(S_t)) - \mathrm{Trace}(\mathbb{E}_{LE}(S_t)) = -\frac{t^4}{12.N}\sum_{i=1}^{N}\mathrm{Trace}\left(L_i^2.\bar{L}_{LE}^2 - (L_i.\bar{L}_{LE})^2\right) + O(t^5).$$

To conclude, we use the following lemma.

LEMMA 4.8. *Let A, $B \in Sym(n)$. Then $Trace(A^2.B^2 - (A.B)^2) \geq 0$. The inequality is strict if and only if A and B do not commute.*

Proof. Let (A_i) (resp., (B_i)) be the column vectors of A (resp., B). Let $\langle , \rangle$ be the usual scalar product. Then we have

$$\begin{cases} \mathrm{Trace}(A^2.B^2) = \sum_{i,j}\langle A_i, A_j\rangle\langle B_i, B_j\rangle, \\ \mathrm{Trace}((A.B)^2) = \sum_{i,j}\langle A_i, B_j\rangle\langle B_i, A_j\rangle. \end{cases}$$

Let us now chose a rotation matrix R that makes A diagonal: $\mathrm{R.A.R^T} = Diag(\lambda_1, \dots, \lambda_n) =: D$. Let us define $C := R.B.R^T$ and use the notation (C_i) and (D_i) for the column vectors of C and D. We have

$$\begin{cases} \mathrm{Trace}(A^2.B^2) = \sum_{i,j}\langle D_i, D_j\rangle\langle C_i, C_j\rangle = \sum_i \lambda_i^2\langle C_i, C_i\rangle, \\ \mathrm{Trace}((A.B)^2) = \sum_{i,j}\langle D_i, C_j\rangle\langle C_i, D_j\rangle = \sum_{i,j}\lambda_i.\lambda_j\langle C_i, C_j\rangle. \end{cases}$$

Then the Cauchy–Schwarz inequality yields

$$\left|\sum_{i,j}\lambda_i.\lambda_j\langle C_i, C_j\rangle\right| \leq \sum_i \lambda_i^2\langle C_i, C_i\rangle,$$

which proves the first point. But the Cauchy–Schwarz inequality is an equality if and only if there is a constant μ such that $D.C = \mu C.D$. But only $\mu = 1$ allows the inequality of the lemma to be an equality. This is equivalent to $C.D = D.C$, which is equivalent in turn to $A.B = B.A$. Hence we have the result. ∎

End of proof of Theorem 4.6. When we apply Lemma 4.8 to the obtained estimation for the trace, we see that for a $t \neq 0$ small enough, the trace of the affine-invariant mean is indeed strictly inferior to the trace of the Log-Euclidean mean whenever the mean logarithm does not commute with all logarithms $\log(S_i)$. ∎

COROLLARY 4.9. *By invariance of the two means with respect to scaling, the strict inequality given in Theorem* 4.6 *is valid in a neighborhood of any SPD matrix of the form λId with $\lambda > 0$.*

COROLLARY 4.10. *When the dimension is equal to* 2*, the Log-Euclidean mean of SPD matrices which are isotropic enough is strictly more anisotropic than their affine-invariant mean when those means do not coincide.*

Proof. In this case, there are only two eigenvalues for each mean. Their products are equal and we have a strict inequality between their sums. Consequently, the largest eigenvalue of the Log-Euclidean mean is strictly larger than the affine-invariant one, and we have the opposite result for the smallest eigenvalue. ∎

4.6. Linear and bilinear interpolation of SPD matrices. Volume elements (or *voxels*) in clinical DT images are often spatially anisotropic. Yet, in many practical situations where DT images are used, it is recommended (see [27]) to work with isotropic voxels to avoid spatial biases. A preliminary resampling step with an adequate interpolation method is therefore important in many cases. Proper interpolation methods are also required to generalize to the SPD case usual registration

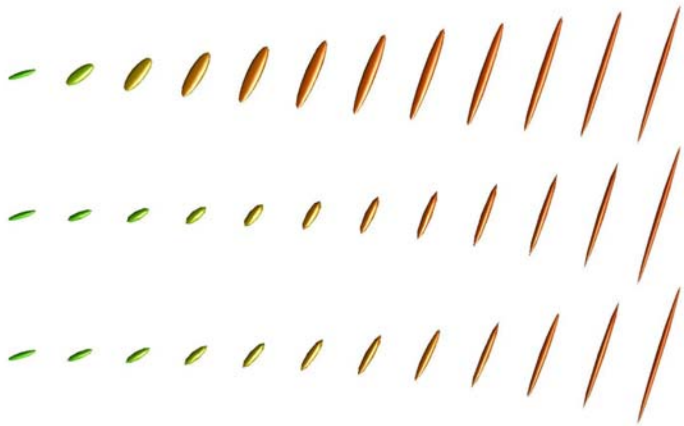

FIG. 4.1. *Linear interpolation of two SPD matrices. Top: linear interpolation on coefficients. Middle: affine-invariant interpolation. Bottom: Log-Euclidean interpolation. The shading of ellipsoids is based on the direction of dominant eigenvectors. Note the characteristic swelling effect observed in the Euclidean case, which is not present in both Riemannian frameworks. Note also that Log-Euclidean means are slightly more anisotropic their affine-invariant counterparts.*

techniques used on scalar or vector images. The framework of Riemannian metrics allows a direct generalization to SPD matrices of classical resampling methods with the use of associated Fréchet means instead of the Euclidean (i.e., arithmetic) mean.

In the Riemannian case, the equivalent of linear interpolation is *geodesic interpolation*. To interpolate between two SPD matrices, intermediate values are taken along the shortest path joining the two matrices. Figure 4.1 presents a typical result of linear interpolation between two SPD matrices. The Euclidean, affine-invariant, and Log-Euclidean results are given. The "swelling effect" is clearly visible in the Euclidean case: the volume of associated ellipsoids is parabolically interpolated and reaches a global maximum between the two extremities! This effect disappears in both Riemannian cases, where volumes are interpolated geometrically. As expected, Log-Euclidean means are a little more anisotropic than their affine-invariant counterparts.

To resample images, bilinear (resp., trilinear) interpolation generalizes in two dimensions (resp., in three dimensions) the linear interpolation and offers an efficient compromise between simplicity and accuracy in the scalar and vector cases. With this technique, the value at any given point is inferred from known values measured at the vertices of a regular grid whose elementary cells are rectangles in two dimensions (resp., right parallelepipeds in three dimensions), which is usually the case with MR images. More precisely, the interpolated value at a given point is given by the weighted mean of the values at the vertices of the current cell. The weights are the *barycentric coordinates* of the current point with respect to the vertices of the current cell.

Figure 4.2 presents the results of the bilinear interpolation of four SPD matrices placed at the extremities of a rectangle. Again, a large swelling effect is present in Euclidean results and not in both Riemannian results, and Log-Euclidean means are slightly more anisotropic than their affine-invariant equivalents. One should note that the computation of the affine-invariant mean here is iterative, since the number of averaged matrices is larger than 2 (we use the Gauss–Newton method described in [12]), whereas the closed form given by 3.8 is used directly in the Log-Euclidean case. This has a large impact on computation times: $0.003s$ (Euclidean), $0.009s$

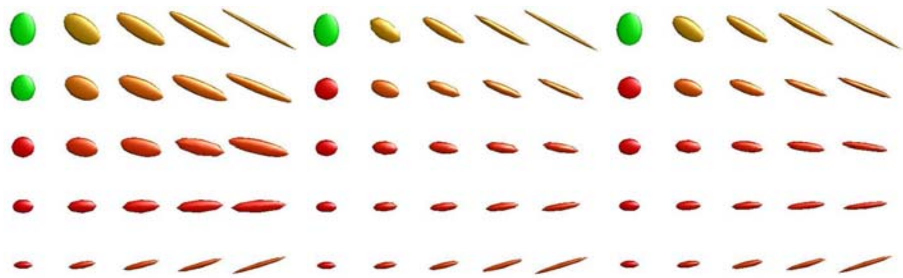

FIG. 4.2. *Bilinear interpolation of four SPD matrices at the corners of a regular grid. Left: Euclidean interpolation. Middle: affine-invariant interpolation. Right: Log-Euclidean interpolation. Again, a characteristic swelling effect is observed in the Euclidean case and not in both Riemannian frameworks. As expected, Log-Euclidean means are slightly more anisotropic than their affine-invariant counterparts.*

(Log-Euclidean), and $1s$ (affine-invariant) for a 5×5 grid on a Pentium M 2 GHz. Computations were carried out with MATLAB, which explains the poor computational performance. Here, Log-Euclidean means were calculated approximately 100 times faster than affine-invariant means because the logarithms of the four interpolated tensors were computed only once, instead of being computing each time a new barycenter is calculated. When only one mean is computed, the typical ratio is closer to 20, since between 15 and 20 iterations are typically needed (for 3×3 SPD matrices) to obtain the affine-invariant mean with a precision of the order of 10^{-12}.

One should note that from a numerical point of view the computation of Log-Euclidean means is not only much faster but also more *stable* than in the affine-invariant case. On synthetic examples, as soon as SPD matrices are quite anisotropic (for instance, with the dominant eigenvalue larger than 500 times the smallest), numerical instabilities appear, essentially due to limited numerical precision (even with double precision). This can greatly complicate the computation of affine-invariant means. On the contrary, the computation of Log-Euclidean means is more stable since the logarithm and exponential are taken only once and thus even very large anisotropies can be dealt with. In applications where very high anisotropies are present, such as the generation of adapted meshes [17], this phenomenon could severely limit the use of affine-invariant means, whereas no such limitation exists in the Log-Euclidean case.

5. Conclusion and perspectives. In this work, we have presented a particularly simple and efficient generalization of the geometric mean to SPD matrices, called Log-Euclidean. It is simply an arithmetic mean in the domain of matrix logarithms. This mean corresponds to a bi-invariant mean in our novel Lie group structure on SPD matrices, or equivalently to a Euclidean mean when this structure is smoothly extended into a vector space by a novel scalar multiplication.

The Log-Euclidean mean is similar to the recently introduced affine-invariant mean, which is another generalization of the geometric mean to SPD matrices. Indeed, the Log-Euclidean mean is similarity invariant, and two means have the same determinant, which is the geometric mean of the determinants of averaged SPD matrices. However, they are not equal: the Log-Euclidean trace is larger when the two means differ. The most striking difference between the two means resides in their computational cost: the Log-Euclidean mean can be calculated approximately 20 times faster than the affine-invariant mean. This property can be crucial in applications where large amounts of data are processed. This is especially the case in medical imaging with DTI and in numerical analysis with the generation of adapted meshes.

We have shown in this work that there are indeed several generalizations of the geometric mean to SPD matrices. Other variants may exist, and we will investigate other possible generalizations in future work. This is important, since situations in applied mathematics, mechanics, medical imaging, etc., where SPD matrices need to be processed, are highly varied. As a consequence, the relevance of each generalization of the geometric mean and of the associated metric framework may depend on the application considered. We have already begun to compare the Log-Euclidean and affine-invariant frameworks in the case of DT-MRI processing [28]. In future work, we will proceed to variability tensors, which we began to use in [15] to model and analyze the variability of brain anatomy.

REFERENCES

[1] S. Lang, *Algebra*, 3rd rev. ed. Grad. Texts in Math., 211 Springer-Verlag, New York, 2002.
[2] M. L. Curtis, *Matrix Groups*, Springer-Verlag, New York, Heidelberg, 1979.
[3] D. Le Bihan, *Diffusion MNR imaging*, Magnetic Resonance Quarterly, 7 (1991), pp. 1–30.
[4] J. Salencon, *Handbook of Continuum Mechanics*, Springer-Verlag, Berlin, 2001.
[5] S. Sternberg, *Lectures on Differential Geometry*, Prentice–Hall, Englewood Cliffs, NJ, 1964.
[6] L. Schwartz, *Analyse Tome* 2: *Calcul Differentiel*, Hermann, Paris, 1997.
[7] N. J. Higham, *Matrix nearness problems and applications*, in Applications of Matrix Theory, M. J. C. Gover and S. Barnett, eds., Oxford University Press, Oxford, UK, 1989, pp. 1–27.
[8] S. Gallot, D. Hulin, and J. Lafontaine, *Riemannian Geometry*, 2nd ed., Springer-Verlag, Berlin, 1990.
[9] R. Godement, *Introduction à la Théorie des Groupes de Lie*, Publications Mathématiques de l'Université Paris VII, Paris, 1982.
[10] M. Wüstner, *A connected lie group equals the square of the exponential image*, J. Lie Theory, 13 (2003), pp. 307–309.
[11] N. Bourbaki, *Elements of Mathematics: Lie Groups and Lie Algebra. Chapters* 1–3, Springer-Verlag, Berlin, 1989.
[12] X. Pennec, P. Fillard, and N. Ayache, *A Riemannian framework for tensor computing*, International J. Computer Vision, 66 (2006), pp. 41–66. A preliminary version appeared as Research Report 5255, INRIA, Sophia-Antipolis, France, 2004.
[13] N. J. Higham, *The scaling and squaring method for the matrix exponential revisited*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 1179–1193.
[14] B. C. Hall, *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*, Grad. Texts in Math., Springer-Verlag, New York, 2003.
[15] P. Fillard, V. Arsigny, X. Pennec, K. M. Hayashi, P. M. Thompson, and N. Ayache, *Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines*, Neuroimage, 34 (2007), pp. 639–650.
[16] T. Broxand, M. Roussonand, R. Deriche, and J. Weickert, *Unsupervised segmentation incorporating colour, texture, and motion*, in Computer Analysis of Images and Patterns N. Petkov and M.A. Westenbere, eds., Lecture Notes in Comput. Sci. 2756, Springer, Berlin, 2003, pp. 353–360.
[17] B. Mohammadi, H. Borouchaki, and P. L. George, *Delaunay mesh generation governed by metric specifications. II. Applications*, Finite Elem. Anal. Des. as, (1997), pp. 85–109.
[18] X. Pennec, *Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements*, J. Math. Imaging Vision, 25 (2006), pp. 127–154. A preliminary version appeared as Research Report RR-5093, INRIA, Sophia-Antipolis, France, 2004.
[19] M. Moakher, *A differential* geometry *approach to the geometric mean of symmetric positive-definite matrices*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 735–747.
[20] C. Feddern, J. Weickert, B. Burgeth, and M. Welk, *Curvature-driven PDE methods for matrix-valued images*, Internat. J. Comput. Vision, 69 (2006), pp. 91–103. Revised version of Tech. Report 104, Department of Mathematics, Saarland University, Saarbrücken, Germany, 2004.
[21] C. Chefd'hotel, D. Tschumperlé, R. Deriche, and O. Faugeras, *Regularizing flows for constrained matrix-valued images*, J. Math. Imaging Vision, 20 (2004), pp. 147–162.
[22] P.T. Fletcher and S.C. Joshi, *Principal geodesic analysis on symmetric spaces: Statistics*

of diffusion tensors., in Proceedings of the CVAMIA and MMBIA Workshops, (Prague, Czech Republic, May 15, 2004), Lecture Notes in Comput. Sci. 3117, Springer, Berlin, 2004, pp. 87–98.

[23] C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras, *Statistics on the manifold of multivariate normal distributions: Theory and application to diffusion tensor MRI processing*, J. Math. Imaging Vision, 25 (2006), pp. 423–444.

%pagebreak

[24] S. Hun Cheng, N. J. Higham, C. S. Kenney, and A. J. Laub, *Approximating the logarithm of a matrix to specified accuracy*, SIAM J. Matrix Anal. Appl., 22 (2001), pp. 1112–1125.

[25] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, *Fast and Simple Computations on Tensors with Log-Euclidean Metrics*, Research Report RR-5584, INRIA, Sophia-Antipolis, France, 2005.

[26] P. Fillard, V. Arsigny, X. Pennec, and N. Ayache, *Clinical DT-MRI estimation, smoothing and fiber tracking with log-Euclidean metrics*, in Proceedings of the Third IEEE International Symposium on Biomedical Imaging (ISBI 2006), Arlington, Virginia, 2006, pp. 786–789.

[27] P. Basser, S. Pajevic, C. Pierpaoli, J. Duda, and A. Aldroubi, *In vivo fiber tractography using DT-MRI data*, Magnetic Resonance in Medicine, 44 (2000), pp. 625–632.

[28] V. Arsigny, P. Fillard, X. Pennec, and N. Ayache, *Log-Euclidean metrics for fast and simple calculus on diffusion tensors*, Magnetic Resonance in Medicine, 56 (2006), pp. 411–421.