Journal of Elasticity (2006) 82: 273–296
DOI: 10.1007/s10659-005-9035-z

On the Averaging of Symmetric Positive-Definite Tensors

MAHER MOAKHER
Laboratory for Mathematical and Numerical Modeling in Engineering Science, National Engineering School at Tunis, ENIT-LAMSIN, B.P. 37, 1002 Tunis-Belvédère, Tunisia.
E-mail: Maher.Moakher@enit.rnu.tn

Received 5 October 2004; in revised form 20 September 2005

Abstract. In this paper we present properly invariant averaging procedures for symmetric positive-definite tensors which are based on different measures of nearness of symmetric positive-definite tensors. These procedures intrinsically account for the positive-definite property of the tensors to be averaged. They are independent of the coordinate system, preserve material symmetries, and more importantly, they are invariant under inversion. The results of these averaging methods are compared with the results of other methods including that proposed by Cowin and Yang (J. of Elasticity **46** (1997) pp. 151–180.) for the case of the elasticity tensor of generalized Hooke's law.

Mathematics Subject Classifications (2000): 74B05, 15A48, 26E60, 47A64.

Key words: symmetric positive-definite tensors, averaging, geometric mean, Kullback–Leibler mean, elasticity tensor.

1. Introduction

It is widely known that positive-definite tensors play important roles in various branches of continuum physics. For instance, it is the positive-definite factor in the polar decomposition frequently used in the analysis of deformations of a continuous medium. Furthermore, linear constitutive laws generally relate two tensors of rank k (with $k = 1$ or 2) via a positive-definite tensor of rank $2k$. For example, Darcy's law of fluid flow in porous media expresses a linear relation between the fluid velocity vector and the gradient of the pore pressure through the second-rank permeability tensor. Similarly, Fourier's law of heat conduction states that the second-rank thermal conductivity tensor relates the heat flux vector to the temperature gradient. More generally, in Fick's law of diffusion the second-rank diffusion tensor relates the flux vector to the concentration gradient. More relevant to the present analysis, in the anisotropic form of Hooke's law of

linear elasticity, the fourth-rank elasticity tensor $\boldsymbol{C}$ relates the stress tensor σ to the infinitesimal strain tensor $\boldsymbol{\epsilon}$, both of which are of second rank, through

$$\sigma = \boldsymbol{C}\boldsymbol{\epsilon}.$$

If a Cartesian orthogonal reference frame is chosen, then Hooke's law can be written in index form as

$$\boldsymbol{\sigma}_{ij} = C_{ijkl}\epsilon_{kl}, \quad 1 \leq i,j,k,l \leq 3.$$

The symmetry of the strain and stress tensors requires that the elasticity tensor satisfies the symmetries

$$C_{ijkl} = C_{jikl} = C_{ijlk}, \quad 1 \leq i,j,k,l \leq 3,$$

known as the minor symmetries. Furthermore, from thermodynamic considerations the elasticity tensor satisfies the symmetry

$$C_{ijkl} = C_{klij}, \quad 1 \leq i,j,k,l \leq 3,$$

known as the major symmetry. We note that further restrictions on the elasticity tensor can be imposed by material symmetries.

In the second-rank tensor notation [16], Hooke's law takes the form

$$\hat{\boldsymbol{\sigma}} = \hat{\boldsymbol{C}}\hat{\boldsymbol{\epsilon}}.$$

Here $\hat{\boldsymbol{\sigma}}$ and $\hat{\boldsymbol{\epsilon}}$ are vectors, and $\hat{\boldsymbol{C}}$ is a second-rank tensor in a six-dimensional space whose components are related to the components of the second-rank tensors σ and $\boldsymbol{\epsilon}$, and the fourth-rank tensor $\boldsymbol{C}$ in the three-dimensional space through the two-way mappings

$$\begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & \sigma_{22} & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & \sigma_{33} \end{pmatrix} \longleftrightarrow \begin{pmatrix} \hat{\sigma}_1 \\ \hat{\sigma}_2 \\ \hat{\sigma}_3 \\ \hat{\sigma}_4 \\ \hat{\sigma}_5 \\ \hat{\sigma}_6 \end{pmatrix} := \begin{pmatrix} \sigma_{11} \\ \sigma_{22} \\ \sigma_{33} \\ \sqrt{2}\sigma_{23} \\ \sqrt{2}\sigma_{13} \\ \sqrt{2}\sigma_{12} \end{pmatrix},$$

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{12} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{13} & \epsilon_{23} & \epsilon_{33} \end{pmatrix} \longleftrightarrow \begin{pmatrix} \hat{\epsilon}_1 \\ \hat{\epsilon}_2 \\ \hat{\epsilon}_3 \\ \hat{\epsilon}_4 \\ \hat{\epsilon}_5 \\ \hat{\epsilon}_6 \end{pmatrix} := \begin{pmatrix} \epsilon_{11} \\ \epsilon_{22} \\ \epsilon_{33} \\ \sqrt{2}\epsilon_{23} \\ \sqrt{2}\epsilon_{13} \\ \sqrt{2}\epsilon_{12} \end{pmatrix},$$

and

$$\left\{ \begin{array}{c} C_{ijkl} \quad 1 \le i,j,k,l \le 3 \\ C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij} \end{array} \right\} \longleftrightarrow \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & \hat{C}_{14} & \hat{C}_{15} & \hat{C}_{16} \\ \hat{C}_{21} & \hat{C}_{22} & \hat{C}_{23} & \hat{C}_{24} & \hat{C}_{25} & \hat{C}_{26} \\ \hat{C}_{31} & \hat{C}_{32} & \hat{C}_{33} & \hat{C}_{34} & \hat{C}_{35} & \hat{C}_{36} \\ \hat{C}_{41} & \hat{C}_{42} & \hat{C}_{43} & \hat{C}_{44} & \hat{C}_{45} & \hat{C}_{46} \\ \hat{C}_{51} & \hat{C}_{52} & \hat{C}_{53} & \hat{C}_{54} & \hat{C}_{55} & \hat{C}_{56} \\ \hat{C}_{61} & \hat{C}_{62} & \hat{C}_{63} & \hat{C}_{64} & \hat{C}_{65} & \hat{C}_{66} \end{pmatrix}$$

$$:= \begin{pmatrix} C_{1111} & C_{1122} & C_{1133} & \sqrt{2}C_{1123} & \sqrt{2}C_{1113} & \sqrt{2}C_{1112} \\ C_{1122} & C_{2222} & C_{2233} & \sqrt{2}C_{2223} & \sqrt{2}C_{2213} & \sqrt{2}C_{2212} \\ C_{1133} & C_{2233} & C_{3333} & \sqrt{2}C_{3323} & \sqrt{2}C_{3313} & \sqrt{2}C_{3312} \\ \sqrt{2}C_{1123} & \sqrt{2}C_{2223} & \sqrt{2}C_{3323} & 2C_{2323} & 2C_{2313} & 2C_{2312} \\ \sqrt{2}C_{1113} & \sqrt{2}C_{2213} & \sqrt{2}C_{3313} & 2C_{2313} & 2C_{1313} & 2C_{1312} \\ \sqrt{2}C_{1112} & \sqrt{2}C_{2212} & \sqrt{2}C_{3312} & 2C_{2312} & 2C_{1312} & 2C_{1212} \end{pmatrix}.$$

The multiplicative factors $\sqrt{2}$ and 2 in the above ensure that the coefficients $\hat{C}_{ij}$ represent the components of a second-rank tensor, see [16] for more details. The second-rank tensor notation is relatively new to the mechanics of solids community which mainly uses the untensorial notation of Voigt [28]. Hence, the elasticity tensor can be expressed either as a fourth-rank tensor in the three-dimensional physical space or as a second-rank tensor in a six-dimensional space. We will call $\hat{\boldsymbol{C}}$ the stiffness tensor and its inverse $\hat{\boldsymbol{S}} = \hat{\boldsymbol{C}}^{-1}$ the compliance tensor.

For later reference we recall the following spectral properties of the stiffness tensor. The eigenvalues and eigenvectors of the stiffness tensor are called the Kelvin eigenvalues and Kelvin eigenvectors, respectively. As the number of distinct eigenvalues is less than or equal to six while for certain material symmetries the number of distinct elastic constants exceeds six, certain eigenvectors must depend on some elastic constants. This dependence is through quantities, called distributors, which are ratios of some elastic constants. Eigenvectors that are independent of the elastic constants are called simple eigenvectors, while eigenvectors that depend on the elastic constants through distributors are called distributor-dependent eigenvectors, or simply distributor eigenvectors [5].

To summarize, all linear constitutive relations mentioned above can be formulated as a linear relation between two vectors via a second-rank tensor. Furthermore, from thermodynamic or energetic considerations this second-rank tensor is symmetric and positive definite.

The coefficients of these tensors with respect to a coordinate system are determined experimentally by a variety of methods. Measurements obtained by different methods for the same material present great variability [7]. One is then confronted with averaging different experimental data sets. A very simple way to

average the experimental data is to average each coefficient separately. However, this method does not take into account that the data are coefficients of positive-definite tensors. Cowin and Yang [6] described another procedure of averaging which consists of two steps. In the first step, the bases of eigenvectors for each matrix of coefficients are averaged. In the second step, the eigenvalues referred to the averaged basis are averaged. In this paper we present other averaging procedures that systematically take into account the positive definiteness of the tensors. These procedures are based on different measures of nearness of symmetric positive-definite tensors. One important feature of these averaging methods that makes them attractive is that they are invariant under inversion. This is very important as some experimental methods give the coefficients of the elasticity (or stiffness) tensor and others give the coefficients of the compliance tensor.

The remainder of this paper is organized as follows. In Section 2, we start by reviewing some geometric properties of the set of symmetric positive-definite matrices and we discuss appropriate metrics to measure the closeness of elements of this set of matrices. We also introduce the Kullback–Leibler divergence. We then use the Euclidean and Riemannian metrics, and the Kullback–Leibler divergence to define properly invariant means for symmetric positive-definite matrices. Applications of these means to averaging anisotropic elastic constant data are presented in Section 3. There the results are compared with those obtained with other methods of averaging.

2. Metrics and Means for Positive-Definite Matrices

In this section we recall some differential-geometric facts of the space of symmetric positive-definite matrices that will be used in the present analysis. Further details can be found in [4, 8, 13, 25]. We also describe means for positive-definite matrices that enjoy certain invariance properties. We only present results that are pertinent to the present analysis. For a more ample analysis and proofs of results not proved here the reader is referred to [18, 19].

Let $\mathbb{M}^{n\times n}$ denote the space of $n \times n$ real matrices and let $GL(n)$ denote the general linear group of all nonsingular matrices in $\mathbb{M}^{n\times n}$. On $\mathbb{M}^{n\times n}$ we have the Frobenius inner product $\langle \boldsymbol{A}, \boldsymbol{B}\rangle_F = \mathrm{tr}(\boldsymbol{A}^T\boldsymbol{B})$, the associated norm $\|\boldsymbol{A}\|_F = [\mathrm{tr}(\boldsymbol{A}^T\boldsymbol{A})]^{1/2}$, and the associated metric

$$\mathrm{d}_F(\boldsymbol{A},\boldsymbol{B}) = \|\boldsymbol{A}-\boldsymbol{B}\|_F. \tag{1}$$

We recall that the exponential map, defined as usual by the absolutely convergent power series $\exp\boldsymbol{A} = \sum_{k=0}^{\infty} \boldsymbol{A}^k/k!$, is a differentiable map from $\mathbb{M}^{n\times n}$ onto $GL(n)$. When a matrix $\boldsymbol{B}$ in $GL(n)$ does not have an eigenvalue in the negative real line, there exists a unique real logarithm, called the principal logarithm and denoted by $\mathrm{Log}\,\boldsymbol{B}$, whose eigenvalues belong to the infinite complex strip

$\{z \in \mathbb{C} : -\pi < \mathrm{Im}(z) < \pi\}$ [11]. We also recall the general fact that for a matrix $\boldsymbol{B}$, such that the principal logarithm is well defined, we have

$$\boldsymbol{A}\ \mathrm{Log}(\boldsymbol{B})\boldsymbol{A}^{-1} = \mathrm{Log}(\boldsymbol{A}\boldsymbol{B}\boldsymbol{A}^{-1}), \tag{2}$$

for every invertible matrix $\boldsymbol{A}$.

2.1. GEOMETRY OF THE SPACE OF SYMMETRIC POSITIVE-DEFINITE MATRICES

The vector space of symmetric matrices in $\mathbb{M}^{n\times n}$ is denoted by $\mathcal{S}(n)$. For $\boldsymbol{P} \in \mathcal{S}(n)$ we say that $\boldsymbol{P}$ is positive semidefinite if the quadratic form $\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{x}$ is non-negative for all $\boldsymbol{x} \in \mathbb{R}^n$. If $\boldsymbol{P}$ is positive semidefinite and invertible we say that $\boldsymbol{P}$ is symmetric positive definite and we write $\boldsymbol{P} > 0$. The subset of $\mathcal{S}(n)$ consisting of all positive-semidefinite matrices is a convex cone whose interior consists of all positive-definite matrices and is denoted by

$$\mathcal{P}(n) := \{\boldsymbol{A} \in \mathcal{S}(n),\ \boldsymbol{A} > 0\}.$$

The exponential of any symmetric matrix is a symmetric positive-definite matrix and the (principal) logarithm of any symmetric positive-definite matrix is a symmetric matrix. Thus the exponential map from $\mathcal{S}(n)$ to $\mathcal{P}(n)$ is one-to-one and onto. It follows that $\mathcal{S}(n)$ provides a parametrization of $\mathcal{P}(n)$ via the exponential map. Another possible parameterization of $\mathcal{P}(n)$ is given by the spectral decomposition, i.e., n positive numbers (representing the eigenvalues) and a special orthogonal matrix (representing the corresponding orthonormal basis of eigenvectors).

2.2. METRICS ON $\mathcal{P}(n)$

We note that $\mathcal{P}(n)$ is a differentiable manifold of dimension $n(n+1)/2$. At any $\boldsymbol{P} \in \mathcal{P}(n)$ the tangent space $T_{\boldsymbol{P}}$ is identified with $\mathcal{S}(n)$. On $T_{\boldsymbol{P}}$ we define the inner product and corresponding norm

$$\langle \boldsymbol{A}, \boldsymbol{B}\rangle_{\boldsymbol{P}} = \mathrm{tr}(\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}^{-1}\boldsymbol{B}), \qquad \|\boldsymbol{A}\|_{\boldsymbol{P}} = \langle \boldsymbol{A}, \boldsymbol{A}\rangle_{\boldsymbol{P}}^{1/2}, \tag{3}$$

that depend on the point $\boldsymbol{P}$ [13]. To measure closeness of two positive-definite matrices $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ one can use the Euclidean distance (1) of the ambient space $\mathbb{M}^{n\times n}$, i.e.,

$$\mathrm{d}_F(\boldsymbol{P}_1, \boldsymbol{P}_2) = \|\boldsymbol{P}_1 - \boldsymbol{P}_2\|_F. \tag{4}$$

However, it might be more appropriate to use the Riemannian distance induced from (3), that is intrinsic to $\mathcal{P}(n)$ and defined by

$$\mathrm{d}_R(\boldsymbol{P}_1, \boldsymbol{P}_2) = \|\mathrm{Log}(\boldsymbol{P}_2^{1/2}\boldsymbol{P}_1^{-1}\boldsymbol{P}_2^{1/2})\|_F = \left[\sum_{i=1}^{n} \ln^2\ \lambda_i\right]^{1/2}, \tag{5}$$

where λ_i, $i = 1, \ldots, n$ are the (positive) eigenvalues of $\boldsymbol{P}_1^{-1}\boldsymbol{P}_2$. The positivity of the λ_i's is a consequence of the similarity of the (in general non-symmetric) matrix $\boldsymbol{P}_1^{-1}\boldsymbol{P}_2$ and the symmetric positive-definite matrix $\boldsymbol{P}_2^{1/2}\boldsymbol{P}_1^{-1}\boldsymbol{P}_2^{1/2}$. It is straightforward to see that the Riemannian distance (5) is invariant under inversion, i.e.,

$$\mathrm{d}_R(\boldsymbol{P}^{-1}, \boldsymbol{Q}^{-1}) = \mathrm{d}_R(\boldsymbol{P}, \boldsymbol{Q}),$$

and is invariant under congruent transformations, i.e.,

$$\mathrm{d}_R(\boldsymbol{P}, \boldsymbol{Q}) = \mathrm{d}_R(\boldsymbol{S}^T\boldsymbol{P}\boldsymbol{S}, \boldsymbol{S}^T\boldsymbol{Q}\boldsymbol{S}), \qquad \text{for all } \boldsymbol{S} \in GL(n).$$

2.3. THE KULLBACK–LEIBLER DIVERGENCE

In information theory, the Kullback–Leibler divergence is used to measure the difference between two probability distributions [12]. We recall that a divergence on a space X is a non-negative function $J(\cdot,\cdot)$ on the Cartesian product space $X \times X$ which is zero only on the diagonal, i.e., $J(x, y) \geq 0$ for all x and y in X and that $J(x, y) = 0$ if and only if $x = y$. The Kullback–Leibler divergence for a pair of zero-mean Gaussian distributions with a pair of (positive-definite) covariance matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$ gives rise to the Kullback–Leibler divergence between the two matrices $\boldsymbol{P}$ and $\boldsymbol{Q}$

$$KL(\boldsymbol{P}, \boldsymbol{Q}) := \mathrm{tr}(\boldsymbol{Q}^{-1}\boldsymbol{P} - \boldsymbol{I}) - \log\,\det(\boldsymbol{Q}^{-1}\boldsymbol{P}). \tag{6}$$

If λ_i, $i = 1, \ldots, n$ denote the (positive) eigenvalues of $\boldsymbol{Q}^{-1}\boldsymbol{P}$ then

$$KL(\boldsymbol{P}, \boldsymbol{Q}) = \sum_{i=1}^{n} (\lambda_i - \log\lambda_\mathrm{i} - 1). \tag{7}$$

Since $x - \log x - 1 \geq 0$ for all $x > 0$ with equality holding only when $x = 1$, it becomes clear from (7) that $KL(\cdot,\cdot)$ defines a divergence on $\mathcal{P}(n)$.

We note here that the Kullback–Leibler divergence (6) is not symmetric with respect to its two arguments. However, it can easily be symmetrized and its symmetric form $KL_s(\boldsymbol{P}, \boldsymbol{Q}) := \frac{1}{2}(KL(\boldsymbol{P}, \boldsymbol{Q}) + KL(\boldsymbol{Q}, \boldsymbol{P}))$ can be expressed as

$$KL_s(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2}\,\mathrm{tr}(\boldsymbol{Q}^{-1}\boldsymbol{P} + \boldsymbol{P}^{-1}\boldsymbol{Q} - 2\boldsymbol{I}), \tag{8}$$

or, in terms of the λ_i's, as

$$KL_s(\boldsymbol{P}, \boldsymbol{Q}) = \frac{1}{2}\sum_{i=1}^{n}\left(\sqrt{\lambda_i} - \frac{1}{\sqrt{\lambda_i}}\right)^2. \tag{9}$$

Now the symmetrized Kullback–Leibler divergence (8) has all the properties of a distance function on $\mathcal{P}(n)$ except that it does not satisfy the triangle inequality. We point out that the symmetrized Kullback–Leibler divergence (8) is invariant under inversion, i.e.,

$$KL_s(\boldsymbol{P}^{-1}, \boldsymbol{Q}^{-1}) = KL_s(\boldsymbol{P}, \boldsymbol{Q}),$$

and is invariant under congruent transformations, i.e.,

$$KL_s(\boldsymbol{P}, \boldsymbol{Q}) = KL_s(\boldsymbol{S}^T\boldsymbol{P}\boldsymbol{S}, \boldsymbol{S}^T\boldsymbol{Q}\boldsymbol{S}), \qquad \text{for all } \boldsymbol{S} \in GL(n).$$

2.4. METRIC-BASED MEANS OF SYMMETRIC POSITIVE-DEFINITE MATRICES

The arithmetic mean of m given symmetric positive-definite matrices $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ is the symmetric positive-definite matrix

$$\mathsf{A}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) := \frac{1}{m} \sum_{k=1}^{m} \boldsymbol{P}_k,$$

which is the unique solution of the minimization problem

$$\min_{\boldsymbol{P} \in \mathcal{P}(n)} \sum_{k=1}^{m} \mathrm{d}_F(\boldsymbol{P}, \boldsymbol{P}_k)^2. \tag{10}$$

As this mean is associated with the Euclidean distance (4), we therefore also call it the Euclidean mean. In analogy with this fact and the variational characterization of various notions of the mean of positive numbers, the author defined in [18] the following metric-based mean for the space of symmetric positive-definite matrices.

DEFINITION 2.1 (Riemannian mean). *The Riemannian mean, i.e., associated with the metric* (5), *of a given set of m symmetric positive-definite matrices* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ *is defined as*

$$\mathsf{G}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) := \arg\min_{\boldsymbol{P} \in \mathcal{P}(n)} \sum_{k=1}^{m} \mathrm{d}_R(\boldsymbol{P}, \boldsymbol{P}_k)^2. \tag{11}$$

It is clear that this definition ensures that the Riemannian mean is invariant under reordering of the matrices to be averaged.

Unlike the minimization problem (10) which can be solved in closed form, the minimization problem (11) for the Riemannian mean leads to a nonlinear matrix equation that, due to the non-commutative nature of matrix multiplications, seems impossible be solved analytically.

PROPOSITION 2.2 ([18]). *The Riemannian mean of m given symmetric positive-definite matrices $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ is the unique symmetric positive-definite solution of the nonlinear matrix equation*

$$\sum_{k=1}^{m} \operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) = \boldsymbol{0}. \tag{12}$$

Except for special cases, there are no known closed-form solution of equation (12). In the case $m = 2$, the Riemannian mean of $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ is given explicitly by any of the six equivalent expressions [18]

$$\begin{aligned} \mathsf{G}(\boldsymbol{P}_1, \boldsymbol{P}_2) &= \boldsymbol{P}_1(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)^{1/2} = \boldsymbol{P}_2(\boldsymbol{P}_2^{-1}\boldsymbol{P}_1)^{1/2} \\ &= (\boldsymbol{P}_2\boldsymbol{P}_1^{-1})^{1/2}\boldsymbol{P}_1 = (\boldsymbol{P}_1\boldsymbol{P}_2^{-1})^{1/2}\boldsymbol{P}_2 \\ &= \boldsymbol{P}_1^{1/2}(\boldsymbol{P}_1^{-1/2}\boldsymbol{P}_2\boldsymbol{P}_1^{-1/2})^{1/2}\boldsymbol{P}_1^{1/2} \\ &= \boldsymbol{P}_2^{1/2}(\boldsymbol{P}_2^{-1/2}\boldsymbol{P}_1\boldsymbol{P}_2^{-1/2})^{1/2}\boldsymbol{P}_2^{1/2}. \end{aligned} \tag{13}$$

The symmetric positive-definite matrix $\mathsf{G}(\boldsymbol{P}_1, \boldsymbol{P}_2)$ corresponds to the notion of geometric mean for a pair of Hermitian operators first introduced by Pusz and Woronowicz [22] and studied thereafter by Trapp [26]. Furthermore, when all matrices $\boldsymbol{P}_k$, $k = 1, \ldots, m$ commute, their Riemannian mean is given by

$$\mathsf{G}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) = (\boldsymbol{P}_1 \cdots \boldsymbol{P}_m)^{1/m} = \boldsymbol{P}_1^{1/m} \cdots \boldsymbol{P}_m^{1/m}.$$

For these reasons, the Riemannian mean is also called the geometric mean. We point out that there are other notions of geometric means. Ando et al. [2] defined the geometric mean of more than two matrices by an iterative process that uses the explicit expressions (13) of the geometric mean of a pair of matrices. For example, the geometric mean of three given matrices is obtained by replacing these matrices by the three geometric means (given explicitly by any of the six equivalent expressions in (13)) of these matrices taken two by two and then iterating until the three matrices obtained are identical. Their geometric mean shares many of the properties of our metric-based geometric mean. However, for more than two matrices these two geometric means are different in general.

2.4.1. *Invariance properties of the geometric mean*

The geometric mean satisfies many invariance properties that make it more attractive than the arithmetic mean. For instance, the invariance under inversion is a property that the arithmetic mean does not satisfy. Some of these properties are given in the following propositions. The third one will be of importance to us in the discussion of material symmetries for the elasticity tensor. The proofs of the first two follow immediately either from the invariance properties of the

Riemannian distance or by simple use of equation (2). We only give a proof of the third one.

PROPOSITION 2.3 (Invariance under congruent transformations). *If* $\boldsymbol{P}$ *is the geometric mean of the set* $\{\boldsymbol{P}_k\}_{1\leq k\leq m}$ *of positive-definite matrices, then* $\boldsymbol{C}\boldsymbol{P}\boldsymbol{C}^T$ *is the geometric mean of the set* $\{\boldsymbol{C}\boldsymbol{P}_k\boldsymbol{C}^T\}_{1\leq k\leq m}$, *for every* $\boldsymbol{C}$ *in* $GL(n)$.

PROPOSITION 2.4 (Invariance under inversion). *If* $\boldsymbol{P}$ *is the geometric mean of the set* $\{\boldsymbol{P}_k\}_{1\leq k\leq m}$ *of positive-definite matrices, then* $\boldsymbol{P}^{-1}$ *is the geometric mean of the set* $\{\boldsymbol{P}_k^{-1}\}_{1\leq k\leq m}$.

PROPOSITION 2.5 (Invariance group). *Let* $\mathcal{A}$ *be a given set of invertible matrices in* $\mathbb{M}^{n\times n}$. *Assume that* $\boldsymbol{P}_1,\ldots,\boldsymbol{P}_m$ *are* m *matrices in* $\mathcal{P}(n)$ *that satisfy*

$$\boldsymbol{P}_k = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{A}^{-1},\quad \forall \boldsymbol{A}\in\mathcal{A},\quad k=1,\ldots m.$$

Then their geometric mean $\boldsymbol{P}$ *also satisfies*

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1},\quad \forall \boldsymbol{A}\in\mathcal{A}.$$

Proof. We note in passing that the set $\mathcal{A}$ is a subset of a group $\mathcal{I}$, such that $\{\boldsymbol{I},-\boldsymbol{I}\}\subseteq\mathcal{I}\subseteq GL(n)$, which we call the invariance group for the matrices $\boldsymbol{P}_1,\ldots,\boldsymbol{P}_m$.

Now recall that the geometric mean of $\boldsymbol{P}_1,\ldots,\boldsymbol{P}_m$ is the unique symmetric positive-definite tensor, $\boldsymbol{P}$, solving the matrix equation

$$\sum_{k=1}^{m}\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}) = \boldsymbol{0}.$$

By left multiplication by $\boldsymbol{A}\in\mathcal{A}$ and right multiplication by $\boldsymbol{A}^{-1}$ of this equation and use of (2) it follows that

$$\sum_{k=1}^{m}\operatorname{Log}(\boldsymbol{A}\boldsymbol{P}_k^{-1}\boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1}) = \boldsymbol{0},$$

which by virtue of the assumption on the $\boldsymbol{P}_k$'s becomes

$$\sum_{k=1}^{m}\operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1}) = \boldsymbol{0}.$$

Then uniqueness of the geometric mean and arbitrariness of $\boldsymbol{A}\in\mathcal{A}$ imply that $\boldsymbol{P}$ must satisfy

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1},\quad \forall \boldsymbol{A}\in\mathcal{A}.$$

□

2.4.2. *Fixed-point algorithm for computing the geometric mean*

As stated earlier, the Riemannian mean of a given set of m symmetric positive-definite matrices $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ is the unique symmetric positive-definite matrix $\boldsymbol{P}$ solution to the nonlinear matrix equation (12) and it looks impossible that this equation can be solved in closed form for $m > 2$. We here describe a fixed-point algorithm to numerically solve the geometric mean of a set of symmetric positive-definite matrices. We mention that other methods such as Newton's method on Riemannian manifolds [24] could also be used for the numerical computation of the geometric mean. However, we found that the fixed-point algorithm described below is simple to implement, does not require a sophisticated machinery, and converges rapidly.

By the invariance of the geometric mean under congruent transformations, we may assume without loss of generality that $\boldsymbol{P}_1 = \boldsymbol{I}$ (this can be achieved by left and right multiplication of each $\boldsymbol{P}_k$ by $\boldsymbol{P}_1^{-1/2}$). In such a case, equation (12) can be written as

$$\operatorname{Log} \boldsymbol{P} = -\sum_{k=2}^{m} \operatorname{Log}(\boldsymbol{P}_k^{-1}\boldsymbol{P}). \tag{14}$$

We then left multiply this equation by $\boldsymbol{P}^{1/2}$ and right multiply it by $\boldsymbol{P}^{-1/2}$ to obtain the equivalent equation

$$\operatorname{Log} \boldsymbol{P} = -\sum_{k=2}^{m} \operatorname{Log}(\boldsymbol{P}^{1/2}\boldsymbol{P}_k^{-1}\boldsymbol{P}^{1/2}), \tag{15}$$

which we exploit to construct the fixed-point algorithm on $\mathcal{S}(n)$ described below. Set $\boldsymbol{S} = \operatorname{Log} \boldsymbol{P}$ so that equation (15) writes

$$\boldsymbol{S} = -\sum_{k=2}^{m} \operatorname{Log}(\exp(\boldsymbol{S}/2)\, \boldsymbol{P}_k^{-1}\, \exp(\boldsymbol{S}/2)). \tag{16}$$

We solve this equation by the following fixed-point iterations on $\mathcal{S}(n)$

$$\boldsymbol{S}^{(0)} = \frac{1}{m}\sum_{k=1}^{m} \operatorname{Log} \boldsymbol{P}_k,$$

$$\boldsymbol{S}^{(l+1)} = \alpha \boldsymbol{S}^{(1)} + (\alpha - 1)\sum_{k=2}^{m} \operatorname{Log}(\exp(\boldsymbol{S}^{(l)}/2)\boldsymbol{P}_k^{-1}\exp(\boldsymbol{S}^{(l)}/2)),\ l \geq 0.$$

The parameter α has to be chosen so that the mapping

$$\boldsymbol{S} \mapsto \alpha \boldsymbol{S} + (\alpha - 1)\sum_{k=2}^{m} \operatorname{Log}(\exp(\boldsymbol{S}/2)\boldsymbol{P}_k^{-1}\exp(\boldsymbol{S}/2)), \tag{17}$$

is a contraction. By considering the scalar case, to have a contraction mapping we must have $\frac{m-1}{m} < \alpha < 1$. Using α of the form $\frac{m+k}{m+k+1}$ for $k > 0$, which does satisfy the above condition, we have found that the fixed-point algorithm does indeed converge in the matrix case.

2.5. KULLBACK–LEIBLER DIVERGENCE-BASED MEAN

In a similar way to the metric-based means defined with the Euclidean and Riemannian metrics we define the Kullback–Leibler divergence-based mean.

DEFINITION 2.6. (Kullback–Leibler mean). *The Kullback–Leibler mean of m given symmetric positive-definite matrices* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ *is defined as*

$$\mathsf{KL}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) := \arg\min_{\boldsymbol{P} \in \mathcal{P}(n)} \sum_{k=1}^{m} KL_s(\boldsymbol{P}_k, \boldsymbol{P}). \tag{18}$$

We note that in the above definition the divergences are not squared.

This minimization problem has a unique solution given explicitly in the following lemma [19].

LEMMA 2.7. *The Kullback–Leibler mean of* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ *is given by the geometric mean of the arithmetic mean of* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ *and the harmonic mean of* $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$*, i.e.,*

$$\mathsf{KL}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) = \mathsf{G}(\mathsf{A}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m), \mathsf{H}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m)),$$

where

$$\mathsf{H}(\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m) := (\mathsf{A}(\boldsymbol{P}_1^{-1}, \ldots, \boldsymbol{P}_m^{-1}))^{-1}.$$

Invariance of the Kullback–Leibler mean (18) under inversion and under congruent transformations follows from the invariance of the Kullback–Leibler divergence (8). Now if $\boldsymbol{P}_1, \ldots, \boldsymbol{P}_m$ are m matrices in $\mathcal{P}(n)$ that satisfy

$$\boldsymbol{P}_k = \boldsymbol{A}\boldsymbol{P}_k\boldsymbol{A}^{-1}, \qquad \forall \boldsymbol{A} \in \mathcal{A}, \quad k = 1, \ldots m,$$

where $\mathcal{A}$ is a given set of invertible matrices in $\mathbb{M}^{n\times n}$, then their Kullback–Leibler mean $\boldsymbol{P}$ also satisfies

$$\boldsymbol{P} = \boldsymbol{A}\boldsymbol{P}\boldsymbol{A}^{-1}, \quad \forall \boldsymbol{A} \in \mathcal{A}.$$

It is interesting to see how the Kullback–Leibler mean, which is associated with a divergence and not a metric, combines the arithmetic and harmonic means via the geometric mean in such a way that makes it invariant under inversion, a property that the arithmetic mean does not satisfy. Furthermore, for $m = 2$ the Kullback–Leibler mean coincides with the geometric mean [19].

3. Application to Averaging the Elasticity Tensor

In this section we use the geometric and Kullback–Leibler means for symmetric positive-definite matrices to average symmetric positive-definite tensors. We also recall an averaging procedure proposed by Cowin and Yang [6]. We discuss certain questions related to material symmetries. We mainly work with the elasticity tensor, but the results apply equally to other symmetric positive-definite tensors such as diffusion tensors, permeability tensors, conductivity tensors, etc. Finally, we illustrate the averaging methods discussed here on previously reported data by Hearmon [7] of elastic constants for some piezoelectric materials.

3.1. METHODS OF AVERAGING SYMMETRIC POSITIVE-DEFINITE TENSORS

Before we proceed in describing the following methods for averaging elasticity tensors we would like to clarify how the means of symmetric positive-definite matrices defined in Section 2 translate into means of symmetric positive-definite tensors. There are two approaches to looking at this matter. First, as we have shown, the geometric and Kullback–Leibler means of symmetric positive-definite matrices are invariant under congruent transformations and in particular orthogonal transformations. This has the implication that under a change of coordinates the given matrices and their means transform according to the tensor rule for coordinate change. Therefore, by the geometric mean (or the Kullback–Leibler mean) of a set of symmetric positive-definite tensors we mean the tensor whose matrix in a coordinate system is the geometric mean (or the Kullback–Leibler mean) of the matrices of the given tensors in this coordinate system. Second, the same analysis that has been done with symmetric positive-definite matrices could be carried out for symmetric positive-definite tensors, and all operations, such as the logarithm, that have been performed on matrices apply to tensors too.

3.1.1. *Method* 1: *Arithmetic mean*

In this method the averaged elasticity tensor is given by the arithmetic mean of the given elasticity tensors. This method is a general procedure for averaging tensors that are not necessarily symmetric positive-definite. Here each coefficient of the stiffness matrix is averaged independently of the other coefficients. If we average the inverses of the given tensors and then invert the resultant tensor we get the harmonic mean which is quite different than the arithmetic mean.

3.1.2. *Method* 2: *Spectral decomposition-based mean*

This method was proposed by Cowin and Yang [6] as an alternative to the classical method of the arithmetic mean which does not take into account the positive-definiteness property of the given tensors. It is based on the spectral

decomposition of symmetric positive-definite tensors. It consists of two steps: in the first step the eigenbases of the given tensors are averaged, and in the second step the eigenvalues referred to the averaged eigenbasis are then averaged. The averaging procedure for the eigenbases is based on the mean associated with the Euclidean distance (1) for the orthogonal matrices of eigenvectors. This mean is in fact given by the polar factor in the polar decomposition of the arithmetic mean of the orthogonal matrices of eigenvectors [6, 18]. We note that a modified version of Cowin and Yang's method that is based on the Riemannian mean of orthogonal matrices [17] could also be used for the averaging of symmetric positive-definite matrices. Once again, using Cowin and Yang's method the inverse of the average tensor of the inverse tensors is different than the average of the tensors.

3.1.3. *Method* 3: *Geometric mean*

This method is based on the geometric mean described in this paper. It intrinsically takes into account the differential-geometrical properties of the space of symmetric positive-definite tensors. It has the advantage of being invariant under inversion. This invariance is very important in practical applications such as the averaging of elasticity tensors that are obtained by several experimental methods either as stiffness coefficients or as compliance coefficients. In general, for more than two tensors this mean cannot be given explicitly and one has to resort to numerical methods, such as the fixed-point algorithm described earlier, to compute it. Beside accounting for the positive-definiteness of the tensors, as we shall see later on, the differential-geometric nature of the Riemannian distance (5) combined with the group property of material symmetries surprisingly make this method preserve material symmetries.

3.1.4. *Method* 4: *Kullback–Leibler mean*

This method is based on the Kullback–Leibler mean and therefore it is invariant under inversion. Although the (symmetrised) Kullback–Leibler divergence does not define a metric on the space of symmetric positive-definite tensors, it yields a notion of mean with many desirable properties. This mean is given in a closed form as the geometric mean of the arithmetic and harmonic means. Owing to the fact that the tensors and their inverses have the same structure when it comes to material symmetries, this method also preserves material symmetries.

3.1.5. *Method* 5: *Harmonic mean*

This methods corresponds to taking the arithmetic mean applied on the inverse tensors and then inverting the result. Like the arithmetic mean, neither is this mean invariant under inversion nor does it take into account the positive-definite character of the tensors.

3.1.6. *Method* 6: *Spectral decomposition-based mean on inverses*

This is Cowin and Yang's method but applied on the inverse tensors. It is not invariant under inversion. The eigenbasis for the average tensor of this method is the same as the eigenbasis for the average of method 2. The inverses of the eigenvalues of the averaged tensor according to this method are given by the arithmetic average of the inverse eigenvalues referred to the averaged eigenbasis.

3.2. REMARKS ON AVERAGING METHODS USED FOR POLYCRYSTALS

The problem of estimating the (effective) elastic properties of polycrystals from those of single crystals is an old one that goes back to Voigt [27] and Reuss [23]. Under appropriate assumptions [14], the elasticity properties can be estimated by averaging the fourth-rank stiffness (or compliance) tensor with a weight function describing the distribution of crystallographic orientations [14, 21]

$$E_{ijkl} = \int_{SO(3)} E^0_{ijkl}(\boldsymbol{R}) f(\boldsymbol{R}) d\boldsymbol{R}.$$

Here E_{ijkl} are the components, with respect to an orthonormal coordinate system, of the averaged stiffness (or compliance) tensor; $E^0_{ijkl}(\boldsymbol{R})$ are the components, with respect to the same coordinate system, of the stiffness (or compliance) tensor for a single crystal with orientation $\boldsymbol{R} \in SO(3)$; $f(\boldsymbol{R})$ is the orientation distribution function (ODF); and $d\boldsymbol{R}$ is the Harr measure of $SO(3)$.

In the mechanics and physics of solid literature, the method of estimating the elastic properties by averaging the stiffness tensor is known as Voigt's average, and that by averaging the compliance tensor is known as Ruess' average. Hill [9] showed that the real average should lie between these two extremes. He suggested that the arithmetic mean between the elasticity tensor obtained by Voigt's method and that by Reuss' method, usually referred to as Hill's average, is a good approximation to the real average tensor. It is noted that none of these three averages satisfies the physical requirement, known as "Lichtenecker's inversion postulate'', that averaging tensors or their inverses should lead to the same material properties [14]. Hill [9] also suggested that, for the rigidity modulus, the geometric mean between the Voigt average and the Reuss average should provide a good approximation to the real average. In this context, we believe that the geometric mean tensor between the Voigt average tensor and the Reuss average tensor should provide a good approximation to the real average tensor. We note that, by construction, this average obeys the inversion invariance.

Based on a suggestion of Aleksandrov and Aizenberg [1], Morawiec [20] and later Matthies and Humbert [14] described an averaging procedure for the

elasticity tensor of textured polycrystals which they termed "geometric mean''. Essentially, they compute the arithmetic (weighted) average of the logarithms of the elasticity tensors then they take the exponential of the resultant tensor. Their "geometric mean'' is different than the geometric mean discussed in this paper, they coincide only when the tensors to be averaged two-by-two commute under multiplication. Their geometric mean does satisfy the invariance under inversion; however, it is not invariant under congruent transformations.

3.3. MATERIAL SYMMETRIES

The material symmetries for an elastic solid material form a multiplicative group, $\mathcal{G}$, which is a subgroup of the full orthogonal group $O(3)$. In linear elasticity theory, the restriction on the fourth-rank elasticity tensor $\boldsymbol{C}$ imposed by material symmetries writes

$$\boldsymbol{Q}(\boldsymbol{C}\boldsymbol{\epsilon})\boldsymbol{Q}^T = \boldsymbol{C}(\boldsymbol{Q}\boldsymbol{\epsilon}\boldsymbol{Q}^T)$$

for every $\boldsymbol{Q}$ in the symmetry group $\mathcal{G}$ of the material and any deformation $\boldsymbol{\epsilon}$. In the six-dimensional tensorial notation this restriction yields that

$$\hat{\boldsymbol{C}} = \hat{\boldsymbol{Q}}^T\hat{\boldsymbol{C}}\hat{\boldsymbol{Q}}, \tag{19}$$

must be satisfied for every six-dimensional orthogonal transformation $\hat{\boldsymbol{Q}}$ representing a three-dimensional orthogonal transformation $\boldsymbol{Q} \in \mathcal{G}$. For a comprehensive account of the representation of three-dimensional orthogonal transformations (including reflections) by six-dimensional orthogonal transformations see [15].

Now the following fundamental question arises. Do the different notions of mean described earlier of a given set of elasticity tensors belonging to a certain class of material symmetry belong to that class? It is clear that the arithmetic, harmonic and the spectral decomposition-based means preserve material symmetries. However, this is not obvious for the geometric and the Kullback–Leibler means. Indeed, the answer to this question for these two means is affirmative and is given in the following corollary.

COROLLARY 3.1. *The geometric and Kullback–Leibler means of a set of elasticity tensors belonging to a class of material symmetry inherit the same symmetry*.

Proof. The proof follows immediately from Proposition 2.5 and the similar statement for the Kullback–Leibler mean. □

In what follows we recall the restrictions on the elasticity tensor for some material symmetries and briefly comment on what coefficients of the stiffness matrix can be given explicitly for the different means.

3.3.1. *Isotropy and cubic symmetry*

In an appropriate coordinate system, the matrix of coefficients of the stiffness tensor for cubic symmetry has the form

$$\hat{\boldsymbol{C}} = \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{12} & 0 & 0 & 0 \\ \hat{C}_{12} & \hat{C}_{11} & \hat{C}_{12} & 0 & 0 & 0 \\ \hat{C}_{12} & \hat{C}_{12} & \hat{C}_{11} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{C}_{44} \end{pmatrix}. \tag{20}$$

(When $\hat{C}_{44} = \hat{C}_{11} - \hat{C}_{12}$, we reduce to the isotropic symmetry.) It has three distinct eigenvalues given by $\hat{C}_{11} + 2\hat{C}_{12}$, $\hat{C}_{11} - \hat{C}_{12}$ and $\hat{C}_{44}$ which are of multiplicity one, two and three, respectively. Furthermore, all eigenvectors are independent of the elasticity constants. Consequently, the tensors to be averaged have the same eigenbasis and therefore for these two cases of material symmetry, the averaged tensors of methods 1 and 2 coincide and the average tensor of method 3 is given in closed form

$$\hat{C}_{11} = \frac{1}{3}\mathfrak{g}(\hat{C}^k_{11} + 2\hat{C}^k_{12}) + \frac{2}{3}\mathfrak{g}(\hat{C}^k_{11} - \hat{C}^k_{12}), \tag{21a}$$

$$\hat{C}_{12} = \frac{1}{3}\mathfrak{g}(\hat{C}^k_{11} + 2\hat{C}^k_{12}) - \frac{1}{3}\mathfrak{g}(\hat{C}^k_{11} - \hat{C}^k_{12}), \tag{21b}$$

$$\hat{C}_{44} = \mathfrak{g}(\hat{C}^k_{44}), \tag{21c}$$

where

$$\mathfrak{g}(X^k) := \sqrt[m]{\prod_{k=1}^{m} X^k}.$$

As for the coefficients of the averaged elastic tensor using method 4 they are given by

$$\hat{C}_{11} = \frac{1}{3}\mathfrak{f}(\hat{C}^k_{11} + 2\hat{C}^k_{12}) + \frac{2}{3}\mathfrak{f}(\hat{C}^k_{11} - \hat{C}^k_{12}), \tag{22a}$$

$$\hat{C}_{12} = \frac{1}{3}\mathfrak{f}(\hat{C}^k_{11} + 2\hat{C}^k_{12}) - \frac{1}{3}\mathfrak{f}(\hat{C}^k_{11} - \hat{C}^k_{12}), \tag{22b}$$

$$\hat{C}_{44} = \mathfrak{f}(\hat{C}^k_{44}), \tag{22c}$$

where

$$\mathfrak{f}(X^k) := \sqrt{\left(\sum_{k=1}^{m} X^k\right)\left(\sum_{k=1}^{m} (X^k)^{-1}\right)^{-1}}.$$

3.3.2. *Tetragonal symmetry*

In an appropriate coordinate system, the matrix of coefficients of the stiffness tensor for tetragonal symmetry has the form

$$\hat{\boldsymbol{C}} = \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & 0 & 0 & 0 \\ \hat{C}_{12} & \hat{C}_{11} & \hat{C}_{13} & 0 & 0 & 0 \\ \hat{C}_{13} & \hat{C}_{13} & \hat{C}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{C}_{66} \end{pmatrix}. \tag{23}$$

The three non-trivial (i.e., of the upper-left 3-by-3 block) eigenvalues are

$$\frac{1}{2}\left[(\hat{C}_{11} + \hat{C}_{33} + \hat{C}_{12}) \pm \sqrt{(\hat{C}_{11} + \hat{C}_{12} - \hat{C}_{33})^2 + 8\hat{C}_{13}^2}\right], \text{ and } \hat{C}_{11} - \hat{C}_{12}.$$

Two of the eigenvectors are distributor eigenvectors. Nevertheless, the average eigenbasis for the eigenbases of eigenvectors can be given explicitly, see [6]. Therefore the average tensor using method 2 can be computed analytically. On the other hand, only the coefficients $\hat{C}_{44}$ and $\hat{C}_{66}$ and the difference $\hat{C}_{11} - \hat{C}_{12}$ can be given explicitly for method 3 and method 4. For method 3 we have

$$\hat{C}_{44} = \mathfrak{g}(\hat{C}^k_{44}), \tag{24a}$$

$$\hat{C}_{66} = \mathfrak{g}(\hat{C}^k_{66}), \tag{24b}$$

$$\hat{C}_{11} - \hat{C}_{12} = \mathfrak{g}(\hat{C}^k_{11} - \hat{C}^k_{12}), \tag{24c}$$

and for method 4 we have

$$\hat{C}_{44} = \mathfrak{f}(\hat{C}^k_{44}), \tag{25a}$$

$$\hat{C}_{66} = \mathfrak{f}(\hat{C}^k_{66}), \tag{25b}$$

$$\hat{C}_{11} - \hat{C}_{12} = \mathfrak{f}(\hat{C}^k_{11} - \hat{C}^k_{12}). \tag{25c}$$

The other coefficients can be determined by the geometric mean or the Kullback–Leibler mean in $\mathcal{P}(2)$ of the matrices

$$\begin{pmatrix} \hat{C}^k_{11} + \hat{C}^k_{12} & \sqrt{2}\hat{C}^k_{13} \\ \sqrt{2}\hat{C}^k_{13} & \hat{C}^k_{33} \end{pmatrix}.$$

3.3.3. *Trigonal symmetry*

In an appropriate coordinate system, the matrix of coefficients of the stiffness tensor for trigonal symmetry has the form

$$\hat{\boldsymbol{C}} = \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & \hat{C}_{14} & 0 & 0 \\ \hat{C}_{12} & \hat{C}_{11} & \hat{C}_{13} & -\hat{C}_{14} & 0 & 0 \\ \hat{C}_{13} & \hat{C}_{13} & \hat{C}_{33} & 0 & 0 & 0 \\ \hat{C}_{14} & -\hat{C}_{14} & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{44} & \sqrt{2}\hat{C}_{14} \\ 0 & 0 & 0 & 0 & \sqrt{2}\hat{C}_{14} & \hat{C}_{11} - \hat{C}_{12} \end{pmatrix}. \tag{26}$$

The eigenvalues are

$$\frac{1}{2}\left[(\hat{C}_{11} + \hat{C}_{33} + \hat{C}_{12}) \pm \sqrt{(\hat{C}_{11} + \hat{C}_{12} - \hat{C}_{33})^2 + 8\hat{C}^2_{13}}\right],$$
$$\frac{1}{2}\left[(\hat{C}_{11} - \hat{C}_{12} + \hat{C}_{44}) \pm \sqrt{(\hat{C}_{11} - \hat{C}_{12} - \hat{C}_{44})^2 + 8\hat{C}^2_{14}}\right],$$

with multiplicity one, one, two and two. All eigenvectors are distributor eigenvectors [6]. Similar to the case of tetragonal symmetry, the average eigenbasis for the eigenbases of eigenvectors can be given explicitly, see [6]. There is no way to obtain the coefficients of the average tensor by method 3 and method 4 but to compute them numerically. However, we can reduce the problem to that of the geometric mean or the Kullback–Leibler mean on a subset of $\mathcal{P}(4)$, i.e., for the matrices

$$\begin{pmatrix} \hat{C}^k_{11} & \hat{C}^k_{12} & \hat{C}^k_{13} & \hat{C}^k_{14} \\ \hat{C}^k_{12} & \hat{C}^k_{11} & \hat{C}^k_{13} & -\hat{C}^k_{14} \\ \hat{C}^k_{13} & \hat{C}^k_{13} & \hat{C}^k_{33} & 0 \\ \hat{C}^k_{14} & -\hat{C}^k_{14} & 0 & \hat{C}^k_{44} \end{pmatrix}.$$

It is rather remarkable that the geometric and Kullback–Leibler means are invariant under this class of material symmetry, but as we have shown this is built in.

3.3.4. *Orthotropic symmetry*

In an appropriate coordinate system, the matrix of coefficients of the stiffness tensor for orthotropic symmetry has the form

$$\hat{\boldsymbol{C}} = \begin{pmatrix} \hat{C}_{11} & \hat{C}_{12} & \hat{C}_{13} & 0 & 0 & 0 \\ \hat{C}_{12} & \hat{C}_{22} & \hat{C}_{23} & 0 & 0 & 0 \\ \hat{C}_{13} & \hat{C}_{23} & \hat{C}_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & \hat{C}_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & \hat{C}_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & \hat{C}_{66} \end{pmatrix}. \tag{27}$$

In this case only the coefficients $\hat{C}_{44}$, $\hat{C}_{55}$ and $\hat{C}_{66}$ can be given explicitly for methods 3 and 4. For method 3 we have

$$\hat{C}_{44} = \mathfrak{g}(\hat{C}^k_{44}), \tag{28a}$$

$$\hat{C}_{55} = \mathfrak{g}(\hat{C}^k_{55}), \tag{28b}$$

$$\hat{C}_{66} = \mathfrak{g}(\hat{C}^k_{66}), \tag{28c}$$

and for method 4 we have

$$\hat{C}_{44} = \mathfrak{f}(\hat{C}^k_{44}), \tag{29a}$$

$$\hat{C}_{55} = \mathfrak{f}(\hat{C}^k_{66}), \tag{29b}$$

$$\hat{C}_{66} = \mathfrak{f}(\hat{C}^k_{66}). \tag{29c}$$

The other coefficients have to be computed.

3.4. AVERAGING HEARMON'S DATA

We illustrate the use of means of symmetric positive-definite tensors discussed above for averaging data for the elasticity constants previously reported by Hearmon [7] and used by Cowin and Yang [6] in their analysis. The data are for piezoelectric crystals with cubic, tetragonal, trigonal and orthotropic symmetries.

3.4.1. *Hearmon's data*

In Table I we give Hearmon's data for the coefficients of the stiffness tensor $\hat{\boldsymbol{C}}$. We refer the reader to [7] for the origin of the data and the experimental method by which they were obtained. We remark that some data are obtained directly as stiffnesses and other are obtained as compliances and then converted to stiffnesses by Hearmon.

Table I. Stiffness constants (units are 10^{11} dynes/cm^2) from Hearmon [7, Extracted from Tables 2–6, pp. 120–122].

$\hat{C}_{11}$	$\hat{C}_{44}/2$	$\hat{C}_{12}$
a) $NaClO_3$ (cubic symmetry)		
4.90	1.17	1.39
4.92	1.19	1.45
5.09	1.18	1.55
4.89	1.17	1.39
4.99	1.17	1.41

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$
b) Ammonium dihydrogen phosphate (tetragonal symmetry)					
7.58	2.96	0.87	0.614	−2.43	1.30
6.17	3.28	0.85	0.592	0.72	1.94
6.89	3.35	0.856	0.595	0.40	1.89
6.77	3.38	0.868	0.608	0.59	1.99

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$
c) Potassium dihydrogen phosphate (tetragonal symmetry)					
7.8	7.7	1.27	0.61	3.23	3.84
6.91	5.56	1.29	0.600	−0.600	1.22
7.08	5.84	1.28	0.633	−0.383	1.55
7.14	5.61	1.27	0.628	−0.49	1.29

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$	$\hat{C}_{14}/\sqrt{2}$
d) Quartz (trigonal symmetry)					
8.51	10.55	5.70	0.70	1.41	−1.69
8.67	10.68	5.79	0.69	1.13	−1.80
8.61	10.71	5.87	0.51	1.05	−1.83
8.69	10.68	5.76	0.69	1.56	−1.74
8.67	10.72	5.80	0.71	1.19	−1.78
8.64	10.46	5.66	0.79	1.36	−1.73

$\hat{C}_{11}$	$\hat{C}_{22}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{55}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$	$\hat{C}_{23}$
e) Rochelle salt (orthotropic symmetry) [a]								
6.82	10.31	9.66	1.04	0.297	0.848	6.83	6.38	7.83
4.25	5.15	6.29	1.25	0.304	0.996	2.96	3.57	3.42
4.06	5.20	6.40	1.22	0.300	0.95	2.56	3.46	3.20
2.55	3.81	3.71	1.34	0.321	0.979	1.41	1.16	1.46

[a] The first and last data sets reported by Hearmon are omitted here. The first because of outliers and the last because it is incomplete.

Table II. Stiffness constants of the average elasticity tensors according to methods 1 to 6.

$\hat{C}_{11}$	$\hat{C}_{44}/2$	$\hat{C}_{12}$
a) $NaClO_3$ (cubic symmetry)		
4.958	1.176	1.438
4.958	1.176	1.438
4.957	1.176	1.437
4.957	1.176	1.437
4.956	1.176	1.437
4.956	1.176	1.437

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$
b) Ammonium dihydrogen phosphate (tetragonal symmetry)					
6.853	3.242	0.861	0.602	−0.180	1.780
6.845	3.257	0.861	0.602	−0.187	1.787
6.721	3.229	0.861	0.602	−0.118	1.760
6.725	3.228	0.861	0.602	−0.126	1.759
6.601	3.214	0.861	0.602	−0.074	1.737
6.600	3.216	0.861	0.602	−0.075	1.738

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$
c) Potassium dihydrogen phosphate (tetragonal symmetry)					
7.232	6.178	1.277	0.618	0.439	1.957
7.210	6.222	1.277	0.618	0.417	1.983
7.030	6.041	1.277	0.618	0.381	1.846
7.035	6.052	1.277	0.618	0.399	1.858
6.844	5.931	1.277	0.617	0.361	1.744
6.835	5.948	1.277	0.617	0.352	1.747

$\hat{C}_{11}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$	$\hat{C}_{14}/\sqrt{2}$
d) Quartz (trigonal symmetry)					
8.632	10.633	5.763	0.682	1.283	−1.762
8.625	10.646	5.764	0.676	1.280	−1.762
8.630	10.629	5.763	0.680	1.284	−1.761
8.630	10.629	5.763	0.680	1.284	−1.761
8.627	10.625	5.762	0.679	1.284	−1.761
8.621	10.639	5.762	0.672	1.280	−1.761

$\hat{C}_{11}$	$\hat{C}_{22}$	$\hat{C}_{33}$	$\hat{C}_{44}/2$	$\hat{C}_{55}/2$	$\hat{C}_{66}/2$	$\hat{C}_{12}$	$\hat{C}_{13}$	$\hat{C}_{23}$
e) Rochelle salt (orthotropic symmetry)								
4.420	6.117	6.515	1.212	0.305	0.943	3.440	3.643	3.978
4.831	6.003	6.218	1.212	0.305	0.943	3.632	3.734	3.774
4.031	5.573	5.978	1.207	0.305	0.941	2.997	3.191	3.461
4.031	5.595	5.971	1.207	0.305	0.941	3.006	3.188	3.468
3.685	5.128	5.485	1.202	0.305	0.940	2.620	2.782	3.015
4.016	4.993	5.152	1.202	0.305	0.940	2.760	2.838	2.785

Table III. Relative Euclidean and Riemannian distances between the indicated means of the elasticity tensors.

	$d_F(\boldsymbol{A},\boldsymbol{H})/\|\boldsymbol{A}\|_F$	$d_F(\boldsymbol{CY},\boldsymbol{YC})/\|\boldsymbol{CY}\|_F$	$d_R(\boldsymbol{A},\boldsymbol{H})$	$d_R(\boldsymbol{CY},\boldsymbol{YC})$
a)	0.0004	0.0004	0.0006	0.0006
b)	0.0355	0.0355	0.0570	0.0565
c)	0.0582	0.0582	0.0879	0.0876
d)	0.0005	0.0005	0.0012	0.0012
e)	0.1949	0.2037	0.2360	0.2425

The first data set reported here is for $NaClO_3$ which is a material with cubic symmetry. The second and third data sets, both of which are for elastic materials with tetragonal symmetry, correspond to ammonium dihydrogen phosphate and potassium dihydrogen phosphate, respectively. The fourth data set is for Quartz which is a material with trigonal symmetry. The last data set is for Rochelle salt which is an orthotropic elastic material.

3.4.2. *Results and discussion*

In Table II we give the averaged coefficients of the stiffness tensor $\hat{\boldsymbol{C}}$ for the different data sets given in Table I. For each material we give the coefficients of the stiffness tensor averaged by the arithmetic mean, Cowin and Yang's mean, geometric mean, Kullback–Leibler mean, harmonic mean and the inverse of Cowin and Yang's mean applied on the compliance tensors.

When the data are clustered, as is the case of (a), (b), (c) and (d), the difference between the coefficients of the average elasticity tensor for the different procedures is relatively small. In contrast, the data for Roshelle salt present a great variability and as a consequence there is a big difference between the different average tensors.

If we denote by $\boldsymbol{A}$, $\boldsymbol{CY}$, $\boldsymbol{G}$, $\boldsymbol{KL}$, $\boldsymbol{H}$ and $\boldsymbol{YC}$ the average elasticity tensor obtained by the averaging methods 1 to 6 then we have found that

$$\boldsymbol{H} < \boldsymbol{G} < \boldsymbol{A}, \quad \text{and} \quad \boldsymbol{KL} < \boldsymbol{A}, \tag{30}$$

where $<$ stands for the Löwner order for symmetric positive-definite tensors, i.e., $\boldsymbol{X} < \boldsymbol{Y}$ means that $\boldsymbol{Y} - \boldsymbol{X} > 0$. The inequality $(30)_1$ (with $<$ replaced by $\leq$ and equality holds only if all tensors to be averaged are equal) is of course similar to the harmonic-arithmetic-geometric mean inequality of positive numbers. Here the statements in (30) hold for the given data and might not be true in general. However, the statements in (30) hold true for the means of a pair of symmetric positive-definite tensors. This fact can be proved for example by using the inequality for scalars, the existence of a common orthogonal tensor that

simultaneously diagonalizes the two given tensors, and the equality of $\boldsymbol{G}$ and $\boldsymbol{KL}$ for the case two tensors.

Finally, in Table III we present the difference in the Frobenius and geometric norms between $\boldsymbol{A}$ and $\boldsymbol{H}$, and $\boldsymbol{CY}$ and $\boldsymbol{YC}$. Once again, these distances are small except for Rochelle salt where these relative measures can be up to 24%. Thus, when some of data are obtained as stiffnesses and the others as compliances averaging methods that are invariant under inversion, such as the geometric mean and the Kullback–Leibler mean, could be the right ones to use.

4. Conclusion and Perspectives

In this paper several methods of averaging anisotropic elastic constant data are discussed and are used to average anisotropic elasticity constants. All of these methods preserve material symmetries but only two of them are invariant under inversion. These are the geometric mean and the Kullback–Leibler mean.

Another application of the averaging procedures will be the determination of effective diffusion (stiffness, permeability, conductivity, etc.) tensors for heterogeneous or composite materials. Two-dimensional analyses of effective tensors for random or chess-board structure indicate that the geometric mean may arise for certain composition structure. In fact, the effective diffusion tensor for a heterogeneous two-dimensional material made of two isotropic materials of different diffusivities with random or chess-board structure is shown to be an isotropic tensor with diffusivity equal to the (scalar) geometric mean of the diffusivities of the composing materials, see for example [10, pp. 37 and 236].

Last but not least, the geometric mean can be used to intrinsically interpolate and smooth diffusion tensor fields from diffusion tensor magnetic resonance imaging (DT-MRI) [3, 19]. In this area there are still many things to explore.

Acknowledgements

The author is grateful to Rachid Deriche at INRIA, France for discussions about the Kullback–Leibler divergence. He also thanks the anonymous reviewer for helpful comments on the earlier version of this paper and for bringing [14] to his attention.

References

1. K. Aleksandrov and L. Aisenberg, A method of calculating the physical constants of polycrystalline materials. *Sov. Phys. Dokl.* **11** (1966) 323–325.
2. T. Ando, C.-K. Li and R. Mathias, Geometric means. *Linear Algebra Appl.* **385** (2004) 305–334.
3. P.G. Batchelor, M. Moakher, D. Atkinson, F. Calamante and A. Connelly, A rigorous framework for diffusion tensor calculus. *Magn. Reson. Med.* **53** (2005) 221–225.

4. R. Bhatia, *Matrix Analysis*. Springer, Berlin Heidelberg New York (1997).
5. S.C. Cowin and M.M. Mehrabadi, On the structure of the linear anisotropic symmetries. *J. Mech. Phys. Solids* **40** (1992) 1459–1472.
6. S.C. Cowin and G. Yang, Averaging anisotropic elastic constant data. *J. Elast.* **46** (1999) 151–180.
7. R.F.S. Hearmon, The elastic constants of piesoelectric crystals. *J. Appl. Phys.* **3** (1952) 120–123.
8. S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*. Academic Press, New York (1978).
9. R. Hill, The elastic properties of crystalline aggregate. *Proc. Phys. Soc. A* **65** (1952) 349–354.
10. V.V. Jikov, S.M. Kozlov and O.A. Oleinik, *Homogenization of Differential Operators and Integral Functionals*. Transl. from Russian by G.A. Yosifian. Springer, Berlin Heidelberg New York (1994).
11. C.S. Kenney and A.J. Laub, Condition estimates for matrix functions. *SIAM J. Matrix Anal. Appl.* **10** (1989) 191–209.
12. S. Kullback and R.A. Leibler, On information and sufficiency. *Ann. Math. Stat.* **22** (1951) 79–86.
13. S. Lang, *Fundamentals of Differential Geometry*. Springer, Berlin Heidelberg New York (1999).
14. S. Matthies and M. Humbert, On the principle of a geometric mean of even-rank symmetric tensors for textured polycrystals. *J. Appl. Crystallogr.* **28** (1995) 254–266.
15. M.M. Mehrabadi, S.C. Cowin and J. Jaric, Six-dimensional orthogonal tensor representation of the rotation about an axis in three dimensions. *Int. J. Solids Struct.* **32** (1995) 439–449.
16. M.M. Mehrabadi and S.C. Cowin, Eigentensors of linear anisotropic elastic materials. *Quart. J. Mech. Appl. Math.* **43**(1) (1990) 15–41.
17. M. Moakher, Means and averaging in the group of rotations. *SIAM J. Matrix Anal. Appl.* **24** (2002) 1–16.
18. M. Moakher, A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.* **26** (2005) 735–747.
19. M. Moakher and P.G. Batchelor, Symmetric positive definite matrices: From geometry to applications and visualization. Chapter 17 In: J. Weickert and H. Hagen (eds.), *Visualization and Image Processing of Tensor Fields*. Springer, Berlin Heidelberg New York (2005).
20. A. Morawiec, Calculation of polycrystal elastic constants from single crystal data. *Phys. Status Solidi, B* **154** (1989) 535–541.
21. P.R. Morris, Averaging fourth-rank tensors with weight functions. *J. Appl. Phys.* **40** (1969) 447–448.
22. W. Pusz and S.L. Woronowicz, Functional calculus for sesquilinear forms and the purification map. *Rep. Math. Phys.* **8** (1975) 159–170.
23. A. Reuss, Berechnung der Fliebgrenze von Mischkristallen auf grund der Plastizitätsbedingung für Einkristalle. *Z. Angew. Math. Mech.* **9** (1929) 49–58.
24. S. Smith, Optimization techniques on Riemannian manifolds. In: *Fields Institute Communications*, Vol. 3, AMS (1994) pp. 113–146.
25. A. Terras, *Harmonic Analysis on Symmetric Spaces and Applications II*. Springer, Berlin Heidelberg New York (1988).
26. G.E. Trapp, Hermitian semidefinite matrix means and related matrix inequalities—an introduction. *Linear and Multilinear Algebra* **16** (1984) 113–123.
27. W. Voigt, Ueber die Beziehung zwischen den beiden Elasticitätconstanten isotroper Körper. *Ann. Phys. (Leipzig)*. **38** (1889) 573–587.
28. W. Voigt, *Lehrbuch der Kristallphysik*. Teubner, Leipzig (1928).