HAL
open science

Geodesic distance on non-singular coherency matrix space in polarization optics

Vincent Devlaminck, Patrick Terrier

To cite this version:

Vincent Devlaminck, Patrick Terrier. Geodesic distance on non-singular coherency matrix space in polarization optics. Journal of the Optical Society of America. A Optics, Image Science, and Vision, 2010, 27 (8), pp.1756-1763. hal-00648826

HAL Id: hal-00648826

https://hal.science/hal-00648826v1

Submitted on 8 Mar 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Geodesic distance on non-singular coherency matrix space in polarization optics

V. Devlaminck & P. Terrier

LAGIS –FRE CNRS 3303 Université Lille 1, Sciences et Technologies – 59655 France
e.mail : vincent.devlaminck@univ-lille1.fr

1 – Introduction.

Since fast and reliable polarimetric equipments are available, many polarimetric imaging systems have been designed for a variety of applications. A wide bibliography about passive imaging polarimetry was given by J.S. Tyo and al. in their review [1] on this topic. Recent examples of the applications of imaging polarimetry include medical [2,3], biological sample analysis [4,5], atmospheric physics [6], or the quality control of manufactured objects [7,8] for instance. From a practical point of view, the imaging polarimetry term is applied for "*Mueller imaging*" or "*Stokes imaging*" as measurement of the Mueller matrix entries or Stokes parameters respectively, attached to each pixel in the image. In this paper we only address the analysis of Stokes imaging and are not interested in corrupted data by external noise such as measurement noise, for instance (for specific papers on this issue see [9,10]) since it is well known of the experimentalists that resulting vectors can be located outside of the Poincaré sphere which contains all vectors with physical meaning. We are dealing with the elements of polarization that depend on a random feature of illuminated objects. Under the hypothesis that external noise has been removed, the resulting Stokes vectors are samples of a random process always located in a manifold known as Poincaré sphere. This set of data has a specific structure and geometry and it may be appropriate to take into account these properties in an estimation process. So, let us first remind the specific structure of the Stokes parameters and relations to coherency matrix.

The coherency matrix $\mathbf{\Phi}$ is the covariance matrix of the components E_i of the complex electric field vector $E(\mathbf{r},t)$, then using the Jones vector formalism such a matrix is given by Eq. (1) :

$$\mathbf{\Phi} = \left\langle \mathbf{E} \otimes \mathbf{E}^{\dagger} \right\rangle = \begin{bmatrix} \left\langle E_1 E_1^{*} \right\rangle & \left\langle E_1 E_2^{*} \right\rangle \\ \left\langle E_2 E_1^{*} \right\rangle & \left\langle E_2 E_2^{*} \right\rangle \end{bmatrix} \tag{1}$$

where † and * stand for a Hermitian and complex conjugate respectively. The brackets denote the time or ensemble average with a hypothesis of ergodicity. As a covariance matrix, $\mathbf{\Phi}$ is a Hermitian nonnegative definite matrix. So, its eigenvalues are real and positive. As a 2x2 complex matrix, $\mathbf{\Phi}$ can be expanded as a linear combination of the four Pauli matrices $\boldsymbol{\sigma}_i$ as given by Eq. (2):

$$\mathbf{\Phi} = \frac{1}{2}\sum_{j=0}^{3} S_j \boldsymbol{\sigma}_j = \frac{1}{2}\begin{bmatrix} S_0 + S_1 & S_2 - iS_3 \\ S_2 + iS_3 & S_0 - S_1 \end{bmatrix} \tag{2}$$

where $S = [\ S_0\ S_1\ S_2\ S_3\ ]^T$ stands for the Stokes vector associated with the light beam. Another complete set of four matrices could be used for the decomposition but the Pauli matrices are the natural ones to use because they underlie the geometry of the Poincaré sphere and provide a fundamental connexion between the Mueller and coherency matrices [11]. Under the linear transformation of the electric field by a matrix $\mathbf{J}$ known as the Jones matrix [12], the coherency matrix $\mathbf{\Phi}_{in}$ of the incident beam transforms into $\mathbf{\Phi}_{out}$ as:

$$\mathbf{\Phi}_{out} = \mathbf{J}\,\mathbf{\Phi}_{in}\,\mathbf{J}^{\dagger} \tag{3}$$

The (*DOP*) *degree of polarization* of the light can be expressed either in terms of the invariants of the coherency matrix [13] (determinant and trace of the matrix) or in terms of the Stokes parameters as:

$$DOP = \left\{ 1 - 4\frac{\det(\mathbf{\Phi})}{\left[\mathrm{Tr}(\mathbf{\Phi})\right]^2} \right\}^{1/2} = \frac{\left(S_1^2 + S_2^2 + S_3^2\right)^{1/2}}{S_0} \tag{4}$$

The semi definite positive condition of coherency matrix involves det(**Φ**) ≥ 0. This condition is equivalent to:

$$S_0^2 - \left(S_1^2 + S_2^2 + S_3^2\right) \geq 0 \tag{5}$$

and then the positive value of the degree of polarization must be smaller than unity. When one eigenvalue of **Φ** is zero, *DOP* =1 and we have a pure state of polarization. When the two eigenvalues of **Φ** are equal, *DOP*=0, and we have a fully unpolarized light. All the other configurations with 0< *DOP* <1 are mixed states as convex combinations of pure states. It is also well known that any polarization state (or coherency matrix) can be uniquely decomposed into a fully polarized part and a fully unpolarized part [14]. The Poincaré sphere representation provides a natural coordinate system:

$$s_1 = \frac{S_1}{S_0}, s_2 = \frac{S_2}{S_0}, s_3 = \frac{S_3}{S_0} \tag{6}$$

to analyze the effect of polarization transformation in which pure states of polarization are mapped onto the surface of the unit radius sphere while partially polarized states lie inside the sphere and unpolarized state lies at the center of the sphere.

The polarimetric effect of a material medium may be regarded as equivalent to that of a system composed of a parallel combination of different pure optical systems. This is true for the polarimetric images of a *natural* scene (this means for classical materials that constitute the most frequently imaged objects). In other words, the Mueller matrix of a system can be considered as a convex sum of Mueller-Jones matrices or also named pure Mueller matrices.[16,17]. The transformation known as Jones matrix when acting on the two components of the transverse vector field, is known as Mueller-Jones matrix when acting on the corresponding Stokes vector. To every Jones matrix there corresponds a Mueller matrix (named Mueller-Jones matrix) but the converse is not true [14-15]. Thus, the state of polarization of the outgoing light beam is a convex sum of pure states and is a mixed state with a degree of polarization strictly smaller than the unity. This is specifically verified with classical examples of polarimetric imaging systems where the temporal and spatial resolution of measurement is not precise enough to guarantee a pure state of polarization for each measured point. So, from an experimentalist point of view, we believe that the pure states of polarization are very unlikely to exist in polarimetric images. In other words, a state of polarization with very small eigenvalues has a very low probability to appear in polarimetric images. Under this hypothesis, the actual set of coherency matrices may be restricted to *HPD*(2) set (the manifold of Hermitian Positive Definite matrices of dimension 2).

On the other hand, from a mathematical point of view, computing on Stokes vectors is difficult due to the severe limitations of the standard Euclidean calculus: The mean of a set of Stokes vectors is a Stokes vector as the result of convex operation over a convex set as already mentioned. However, more complex operations like gradient descent or Principal Component Analysis (PCA) that are classical in standard Euclidean space, are not stable in Stokes vector space. One can quickly reach the boundaries of the space.

We will demonstrate the Riemannian metric is a natural way to lead to a solution to these constraints and its introduction on *HPD*(2) may be directly related to the Jones calculus. After a reminder of relations associated with the definition of a Riemannian metric on *HPD*(2), we derive the expression of the distance and introduce the notion of mean value in this particular metric space. We address the question of why use other distances than the Euclidean one and what are the appropriate data to analyze with these distances. Eventually, applications on simulated data and real data of a geodesic version of the classical K-means clustering algorithm are presented.

2 – Riemannian metric for polarization space.

2.1 Mathematical background.

We consider a Stokes vector **S.** Its corresponding coherency matrix **Φ** was given by Eq. (2). If both the eigenvalues of **Φ** are non zero, **Φ** is an element of *HPD*(2), the manifold of Hermitian Positive Definite matrices of dimension 2. A smooth action group of a Lie group *G* on a manifold *M* is defined ([18], p.90) as a mapping $\varphi : G \times M \rightarrow M$ satisfying two conditions:

$$\begin{aligned} &\forall x \in M, \quad \varphi(e,x) = x \\ &\forall (g_1,g_2) \in G \times G, \ \forall x \in M, \quad \varphi(g_1, \varphi(g_2,x)) = \varphi(g_1 g_2, x) \end{aligned} \tag{7}$$

where *e* is the identity element of *G*. We choose to use italic characters to denote the elements of a group and bold characters for their matrix representation. Demonstrating (see Appendix A) that the mapping associated with Eq. (3), defines a smooth action group of $GL(2,\mathbb{C})$ on *HPD*(2), is straightforward. $GL(2,\mathbb{C})$ denoted the group of all non singular 2x2 matrices with complex entries. In order to construct a Riemannian metric on the manifold *HPD*(2), we use the following theorem from [18] p.357. It is worth noting that a similar demonstration is proposed in [19] but for the set of real symmetric definite matrices.

Theorem 1

"*Let G be a Lie group acting transitively on a manifold M. Thus M as a Riemannian metric such that the transformation determined by each element of G is an isometry if the isotropy group H of a point p ∈ M is a connected compact Lie subgroup of G*".

The importance of the Riemannian manifolds derives from the fact that it makes the tangent space $T_p(M)$ at each point *p* into a Euclidean space, with inner product denoted by $< X_p , Y_p >_p$ on $T_p(M)$. This enables us to define *geodesics* and the length of curves on *M*. Let *p* be a point of *M*, and X_p a tangent vector at *p*, there is a unique geodesic $\gamma_{(p,X)}(t)$ on *M* with initial point $\gamma_{(p,X)}(0) = p$ and tangent vector $\gamma'_{(p,X)}(0) = X_p$. $\gamma_{(p,X)}(1)$, the point on the geodesic at which the parameter takes the value 1, is the *exponential mapping* noted $\mathrm{Exp}(X_p)$ or $\mathrm{Exp}_p(X)$. Since $|d\gamma(t)/dt|$ is constant along a geodesic, the length L from $\gamma_{(p,X)}(0)$ to $\gamma_{(p,X)}(1)$ is:

$$L = \int_0^1 \left\| \frac{d\gamma(t)}{dt} \right\| dt = \int_0^1 \|X_p\| \, dt = \|X_p\| \tag{8}$$

and $\mathrm{Exp}(X_p)$ is the point on the geodesic $\gamma_{(p,X)}(t)$ determined by X_p whose distance from *p* along the geodesic is the length of X_p.

We demonstrate in appendix B, $GL(2,\mathbb{C})$ acts transitively on *HPD*(2) and the *isotropy subgroup* of **Id** under the mapping φ, is *U*(2), the group of unitary matrices of dimension 2. As *U*(2) is a connected compact Lie group, *HPD*(2) has a metric that is invariant under the $GL(2,\mathbb{C})$ action defined by φ. The inverse mapping of Exponential mapping is defined by $X = \mathrm{Log}_p(x)$ and named *Log mapping*.

With $\mathbf{y} = \varphi(g^{-1}, x) = \mathbf{g}^{-1}\,\mathbf{x}\,\mathbf{g}^{-\dagger}$:

$$\mathbf{y} = \varphi\left[g^{-1}, \gamma_{(p,X)}(1)\right] = \gamma_{(Id, g^{-1} X g^{-\dagger})}(1) = \exp\left(\mathbf{g}^{-1}\mathbf{X}\mathbf{g}^{-\dagger}\right) \tag{9}$$

$$\mathbf{y} = \exp\left(\mathbf{g}^{-1}\mathbf{X}\mathbf{g}^{-\dagger}\right) \;\Rightarrow\; \mathbf{X} = \mathrm{Log}_p(x) = \mathbf{g}\log(\mathbf{y})\mathbf{g}^{\dagger} \tag{10}$$

where *log* stands for the classical logarithm of matrix. So, it is possible to derive a practical expression of the Riemannian distance between two points p and x of $HPD(2)$. With X, defined as $X = \mathrm{Log}_p(x)$, $p = g\,g^{\dagger}$ and $\mathbf{y} = \varphi(g^{-1}, x) = \mathbf{g}^{-1}\,\mathbf{x}\,\mathbf{g}^{-\dagger}$ (from Eqs. (8)(10) and (B2)):

$$\begin{aligned} \mathrm{d}^2(p,x) &= \langle X, X\rangle_p = \left\|\mathrm{Log}_p(x)\right\| \\ \mathrm{d}^2(p,x) &= \left\langle \mathrm{d}\varphi\left(g^{-1}, X_p\right), \mathrm{d}\varphi\left(g^{-1}, X_p\right)\right\rangle_{Id} = \left\langle g^{-1}X_p g^{-\dagger}, g^{-1}X_p g^{-\dagger}\right\rangle_{Id} = \langle \log(\mathbf{y}), \log(\mathbf{y})\rangle_{Id} \end{aligned} \tag{11}$$

Then using the classical Euclidean matrix norm on the tangent space:

$$\mathrm{d}(p,x) = \left[\mathrm{Tr}\left(\log(\mathbf{y})^2\right)\right]^{1/2} \tag{12}$$

where Tr (**A**) denotes the trace of matrix **A**.

From an algorithmic point of view, the computation of the distance between **p** and **x** is based on the diagonalization of both these matrices and we have (see Appendix B) with $\mathbf{g}=\mathbf{p}^{1/2}$:

$$\begin{aligned} \mathrm{Exp}_p(X) &= \mathbf{x} = \mathbf{p}^{1/2}\exp\left[\mathbf{p}^{-1/2}\mathbf{X}\,(\mathbf{p}^{-1/2})^{\dagger}\right](\mathbf{p}^{1/2})^{\dagger} \\ \mathrm{Log}_p(x) &= \mathbf{X} = \mathbf{p}^{1/2}\log\left[\mathbf{p}^{-1/2}\mathbf{x}\,(\mathbf{p}^{-1/2})^{\dagger}\right](\mathbf{p}^{1/2})^{\dagger} \end{aligned} \tag{13}$$

With Eq. (13), demonstrating that the distance can also be computed as the sum of the square logarithm of the eigenvalues of $p^{-1}x$ (generalized eigenvalues of p and x), is straightforward.

2.2 Properties and physical interpretation of this distance.

We now address the question of why use other distances than the Euclidean one and what are the appropriate data to analyze with these distances. By selecting the metric, we define the topology of the set. Therefore, if more than one metric is admissible, selecting among them must be based on additional information and constraints derived from the topology of the space or from the physical nature of the quantity.

When making noisy measurements of Stokes parameters or coherency matrix, the accuracy of the estimation is increased by performing repetitive measurements, under the assumption of constant quantities over time. So we deal with different realizations of a Stokes vector (or coherency matrix) in the same pixel and we may expect that temporal interpolation will preserve the intensity. Euclidean metric is thus well-adapted since the trace of the coherency matrix is linearly interpolated with this metric. This is the well known principle of summation of waves. The coherency matrix as a physical quantity has a non negative determinant. However, noise or artifacts can result in Stokes parameter measurements leading to negative values of the coherency matrix determinant. The positiveness requirement does not seem plausible in this context, a Euclidean space and its classical associated distance seem to be appropriate. With the summation of waves, the energy conservation principle is obvious. The consequence on the distance properties are trace conservation.

But if two pixels measure the Stokes parameters or coherency matrix in different points along the same object, this latter principal of intensity preservation is not so obvious. To illustrate this statement, we will first examine an elementary example. This example is just chosen as an illustration since it is based on fully polarized Stokes vectors corresponding to coherency matrices that are not elements of $HPD(2)$. Nevertheless extending the conclusions to the elements of $HPD(2)$, is straightforward.

We consider the measurement of Stokes parameters of a horizontally linear polarized light passing through a perfect linear polarizer with a rotating axis of polarization. The resulting Stokes vector has the following expression :

$$\left[\frac{S_0}{2}\left[1+\cos(2\theta)\right] \quad \frac{S_0}{2}\left[1+\cos(2\theta)\right]\cos(2\theta) \quad \frac{S_0}{2}\left[1+\cos(2\theta)\right]\sin(2\theta) \quad 0\right]^T \tag{14}$$

where θ stands for the angle between the rotating axis of the polarizer and the horizontal polarization. The point locus is represented by the blue curve in Fig. 1. Assuming a uniform distribution of θ parameter between - π/2 and +π/2 for example, gives a Euclidean mean Stokes vector S_{mean} = [S_0 /2 , S_0 /4, 0, 0] represented by a blue star point in Fig. 1. When using the Euclidean distance for calculating the mean value, the resulting mean vector outgoing of a perfect linear polarizer is then a partially polarized one. However, a more physical meaning solution seems to be possible when defining the mean Stokes vector as the outgoing vector for the mean position of the polarizer. In this case, the mean vector is the outgoing vector for θ=0 and we have S_{mean} = [S_0, S_0, 0, 0] represented by a red star point in Fig. 1. This is also, the mean vector corresponding to the solution derived from the geodesic distance applied to the manifold described by the Eq. (14) with variables S_0 and θ. Then the Euclidean mean Stokes vector does not always seem to be the best solution.

In another part, it is worth noticing that the same input Stokes vector passing through a rotator with a rotation angle θ, produces an outgoing vector with the following expression [S_0 , S_0 cos(2θ), - S_0 sin(2θ), 0] and the same geodesic mean Stokes vector. The point locus is represented by the red curve in Fig. 1 with S_0 =1. In the Poincaré representation, both these outgoing vectors describe the same set of points. The only way to discriminate them is to take into account their intensity components. But, due to the severe limitations of the standard Euclidean calculus, problems appear when the intensity component is explicitly taken into account since it is treated independently from Eq. (5) constraint. As a consequence, methods like clustering algorithms, often work only on reduced components or on physical quantities derived from states of polarization of the wave (angles of polarization, *DOP,..*) but lose a part of the information. This problem does not happen with the geodesic approach.

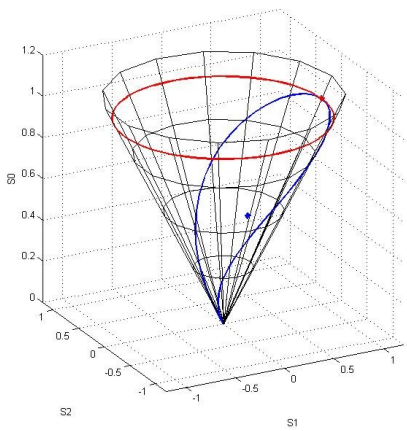

Fig. 1: 3D representations (S3=0) of point locus of outgoing Stokes vectors from a polarizer (blue curve) and from a rotator (red curve) for a fully polarized input vector with unitary intensity. The blue star point is the Euclidean mean vector and the red star point is the geodesic mean vector.

The third point is that we have a context dependent solution [21]. Each solution assumes an underlying manifold where the data are located and a related geodesic distance. For instance, in the previous examples, data are only authorized to be located on the blue or red line (see Fig. 1) but not in the full space of the light cone or in the full Euclidean space. Since Stokes vectors or coherency matrices can have a maximum of four degrees of freedom, the dimension of the related manifolds takes values from 1 to 4. It is possible to capture the underlying manifold structure of the data and determine its dimension by using a manifold learning technique called Isomap [21-23. It is obvious that the proposed geodesic distance is related to data living on a manifold structure of dimension 4, since there is no specific restriction on the *HPD(2)* manifold we consider.

Having formulated the Riemannian distance of polarimetric space and established when we can use this distance, we now apply this definition for computing statistics in this space.

2.3 Average of coherency matrices.

The *intrinsic mean* μ of a random variable x in an arbitrary metric space was defined by Fréchet [24] as the points that minimize the variance:

$$\sigma_x^2(p) = E\left[d(p,x)^2 \right] \tag{15}$$

where d is geodesic distance on the Riemannian manifold *M* and E(x) denotes the expectation of the random variable x. This intrinsic mean is thus obtained as the result of a minimization process :

$$\mu = \arg\min_{p \in M} \left(E\left[d(p,x)^2 \right] \right) \tag{16}$$

Pennec [25] describes an iterative algorithm based on a gradient descent technique to compute this minimum (the question of existence and uniqueness of the mean is not treated in this paper but see [25] [26] for details on this topic). With the input $x = x_1,..., x_n$, output μ is the result of the gradient descent term:

$$\mu_{i+1} = \mathrm{Exp}_{\mu_i}\left(\frac{1}{n} \sum_{j=1}^{n} \mathrm{Log}_{\mu_i}(x_j) \right) \tag{17}$$

It is worth noticing that this definition of the intrinsic mean value gives a test condition to experimentally characterize non singular Mueller-Jones matrices. Since the mapping associated with Eq. (3), defines a smooth action group of *GL*(2,ℂ) on *HPD*(2), the transformation determined by each element of *GL(2,ℂ)* is an isometry (from Theorem 1). Thus, the Riemannian distance of Eq. (12) between two coherency matrices (two Stokes vectors respectively) is not modified by the action of a Jones matrix (Mueller-Jones matrix respectively). According to the definition of the mean value previously introduced, this property can be applied to the average of coherency matrices. Given a set of input Stokes vectors S_{in} and S_{out} the set of input vectors multiplied by a Mueller matrix **M, i**f this matrix is a Mueller-Jones one, the mean value of S_{out} is the product of mean value of S_{in} by **M**.

Another consequence of the distance properties is related to the variance defined by Eq. (15). The variance of a random coherency matrix is not affected by the action of a non-singular deterministic Jones matrix. In other words, this variance is independent from the particular polarization properties of the mean value when the congruency transformation defined by Eq. (3) is acting on a random coherency matrix. It is worth noticing that the corresponding property exists for the Euclidean space but with the translation operation.

3 – Experimentation on simulated and real data.

It is important to understand the structure of the space induced by metric. In the Euclidean case, the Poincaré sphere is a part of $\mathbb{R}^3$ where the sphere is embedded (or the hyper-cone is part of $\mathbb{R}^4$ for the Stokes vectors). Geodesics can also be defined in the Euclidean space. These geodesics as the shortest paths between two points in the Euclidean space, are the well known straight lines in this case. As elements of $\mathbb{R}^3$, these straight lines are not constrained to stay inside the Poincaré sphere, the physical solution field. In the Riemannian case, all matrices are by nature Hermitian Positive Definite. The Riemannian distance of Eq. (12) involves a non linear structure of the associated space. This property can be verified with the example of Fig. 2. The distance between a totally unpolarized Stokes vector and a Stokes vector with variable *DOP* is depicted. Intensity value is fixed to 2. For the

extrinsic Euclidean distance on the Poincaré sphere, the value of the distance between these Stokes vectors is given by the value of the *DOP* (Fig. 2, dotted line). For the intrinsic geodesic distance, a numerical calculus from Eq. (12) gives the values of the distance (Fig. 2, solid line). Since one of both matrices is proportional to identity matrix, the generalized eigenvalues are the eigenvalues of the other matrix and in this case, an analytical expression of the geodesic distance as function of eigenvalues of the other matrix, can be derived. These eigenvalues are functions of *DOP* and of the intensity S_0 of the wave [11]. A straightforward calculus gives:

$$\mathrm{d}(\mathrm{Id}, x) = \left[\log^2\left[\frac{S_0(1+DOP)}{2}\right] + \log^2\left[\frac{S_0(1-DOP)}{2}\right]\right]^{1/2} \tag{18}$$

showing an expression of the distance as a function of intensity and degree of polarization. It is worth noticing that the result is obtained with one of the matrices equal to the identity one, and is not a general property of the proposed geodesic distance. If both matrices are different from the identity one, the distance depends on the intensity and polarization state and not only of degree of polarization. Non linear effects induced by the expression of Eq. (18) are reported by Fig. 2 where intensity value is fixed to 2.

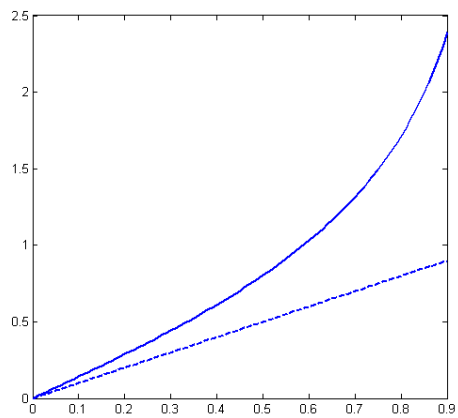

Fig. 2 – Euclidean distance (dotted line) and geodesic distance (solid line) as function of *DOP* with S_0 =2

As consequences of non linear effects on statistics, random variables with the same geodesic variance as defined by Eq. (15) should have different Euclidean representations according to the mean state of polarization. So, the notion of polarimetric contrast may be revisited. This statement can be illustrated by the example of normal distributions of coherency matrices on the *HPD(2)* manifold.

3.1 Simulated data with a normal law.

We define Gaussian distributions on Lie group following Grenander [27], as a solution to the heat equation defined in the local coordinates of the Lie group. To construct the corresponding normal law, we use the approach proposed by Langlet & al. [28] based on the works of Pennec [25] The property that is taken for granted is the maximization of the entropy knowing the mean and the covariance matrix. These authors show that such a method can provide computationally tractable approximations for any manifold in case of small variances.

We generate a set of random coherency matrices so that they follow a generalized normal law with a mean coherency matrix $\mathbf{\Phi}$ and a covariance matrix $\mathbf{\Lambda}$. As the coherency matrices have only four degrees of freedom (the four Stokes vector components for instance) the covariance matrix $\mathbf{\Lambda}$ can be defined by the covariance matrix of the associated Stokes vector components. See Appendix C for a description of the corresponding algorithm that we adapt from [28] to the case of *HPD(2)* manifold. Fig. 3 shows the points locus in the Poincaré sphere representation of samples generated by the algorithm with different configurations of mean coherency matrix and the state of polarization of the mean point.

The value of the variance is constant and fixed to 0.005. Samples in Fig.3-a correspond to a mean coherency matrix associated to the Stokes vector [1, DOP, 0, 0] with DOP = 0.1 (blue), 0.6 (red) and 0.9 (green). Samples in Fig. 3-b correspond to a mean coherency matrix associated to the Stokes vector [0.3742/DOP 0.1, 0.2, 0.3] with DOP = 0.1 (blue), 0.6 (red) and 0.9 (green). In both figures, points for DOP = 0.9 are represented after a sign inversion in order to have a better representation. Although *DOP* has the same values in the both configurations, the set of generated states of polarization highly depends on intensity values.

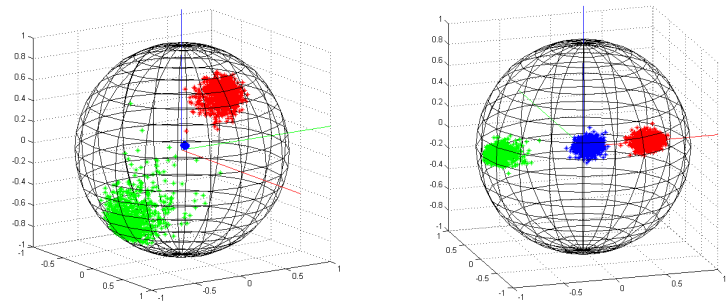

Fig. 3 : Examples of Poincaré sphere representation of samples with generalized Gaussian distribution for a variance fixed to 0.005. The mean point has a *DOP* of 0.1 (blue), 0.6 (red) and 0.9 (green). (a) : Intensity of the mean point is fixed to 1. (b): intensity of the mean point is variable.

The value of the variance is constant and fixed to 0.005. Samples in Fig.3-a correspond to a mean coherency matrix associated to the Stokes vector [1, DOP, 0, 0] with DOP = 0.1 (blue), 0.6 (red) and 0.9 (green). Samples in Fig. 3-b correspond to a mean coherency matrix associated to the Stokes vector [0.3742/DOP 0.1, 0.2, 0.3] with DOP = 0.1 (blue), 0.6 (red) and 0.9 (green). In both figures, the points for DOP = 0.9 are represented after a sign inversion in order to have a better representation. Although *DOP* has the same values in both configurations, the set of generated states of polarization highly depends on intensity values.

3.2 Mean value estimation.

Based on these simulated data, we perform an estimation of the mean coherency matrix using the algorithm depicted in section 2.3. Rather than the estimated values of matrices, which are not easy to analyze, we prefer to plot the computed *DOP* and polarization angle of the wave associated with these estimated mean coherency matrices. Fig. 4-a and 4-b show these estimated values (green star points) of *DOP* and polarization angle respectively. The mean coherency matrices to estimate, are related to the Stokes vectors [1, DOP, 0, 0] for DOP values between 0.05 and 0.95.The theoretical values are plotted as a blue line. The estimation with the classical Euclidean distance are plotted as a red line. Fig. 5-a and 5-b show the same quantities but for mean coherency matrices estimation related to the Stokes vectors [0.3742/DOP 0.1, 0.2, 0.3]. These results show the ability of our method to correctly estimate the coherency matrix and the difficulties of Euclidean estimation for points closed to the boundaries of the manifold (when the *DOP* is increasing).

3.3 Real data

As an illustration of advantages of geodesic distance to analyze real data, we address the K-means clustering technique [29]. K-means clustering partitions observations into K clusters in which each observation belongs to the cluster with the nearest mean. This algorithm uses a two-phase iterative process to minimize the sum of point-to-centroid distances. In a first phase each iteration consists of reassigning points to their nearest cluster centroid. Cluster centroids are recomputed. In the second phase, points are individually reassigned if doing so will reduce the sum of distances. Cluster centroids

are recomputed after each reassignment. So this technique may be based either on the classical Euclidean distance or on the proposed geodesic distance. From a practical point of view, we define the *Geodesic K-means* algorithm as a K-means algorithm on the tangent space. It means that the samples are first transformed by the Log mapping as an element of the tangent space. A classical K-means algorithm is applied on this linear space and the results are transformed by the Exponential mapping as elements of the original space. Obviously, this is only a first order approximation (linear approximation) of a full geodesic K-means algorithm. The error is minimal around 0 and as it is always possible to define a mapping that is a tangent vector at **Id** (see previous paragraphs), this mapping assures that the error is minimized. These clustering algorithms are tested on real data. Fig. 6 shows an example of "Stokes imaging" as a measurement of the four entries of Stokes vectors attached to each pixel in the image. This is an image of a thin dielectric deposit (text) on a metallic object (the text was written with a lacquer for hair). In the experimental configuration, the polarimeter is fixed in the vertical reference direction, a linear +45° polarized light source ($S = [S_0, 0, S_0, 0]^T$) is illuminating the objects. A complete description of the measurement system can be found in [7]. Following the previous remarks of section 2.2 about the importance of the intensity component in clustering process, we applied the proposed geodesic approach on coherency matrices associated to the Stokes parameters of Fig. 6.

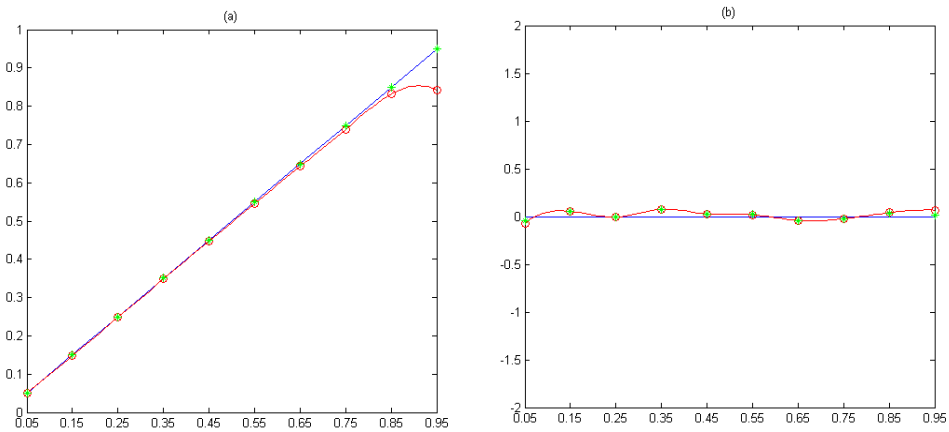

Fig. 4 : Estimated values (green star points) of *DOP* (a) and polarization angle (b) for the Stokes vector [1, DOP, 0, 0] as function of DOP values The theoretical values are plotted as a blue line. Estimation with the classical Euclidean distance are plotted as a red line.

Fig. 7 shows the resulting clusters for both methods where the clusters number is fixed to 3. The pixels position belonging to these 3 Euclidean or geodesic clusters are depicted in Fig. 8 using a three grey levels scale. These images clearly demonstrate the advantages of the geodesic approach

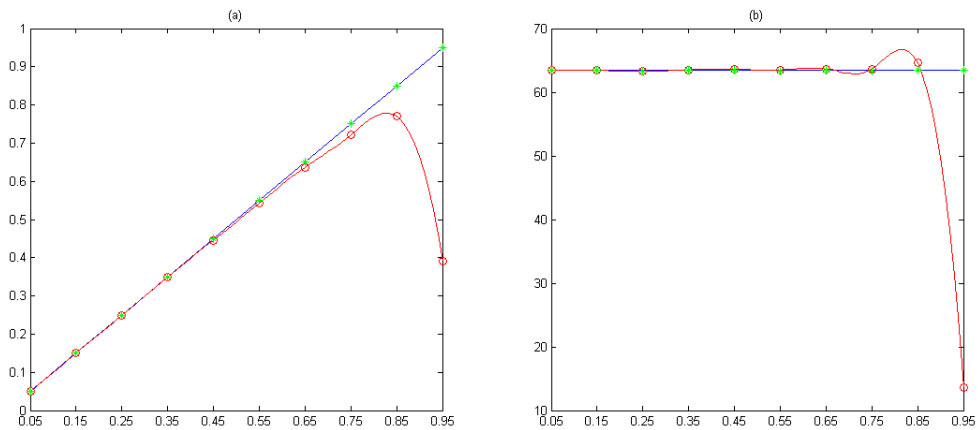

Fig. 5 : Estimated values (green star points) of *DOP* (a) and polarization angle (b) for the Stokes vector [0.3742/DOP 0.1, 0.2, 0.3] as function of DOP values The theoretical values are plotted as a blue line. Estimation with the classical Euclidean distance are plotted as a red line.

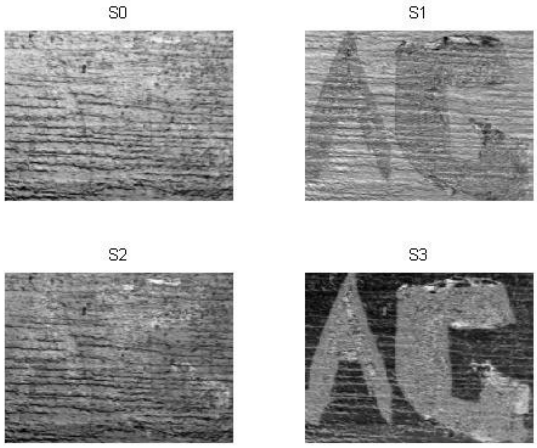

Fig. 6: Images of the Stokes parameters of a thin dielectric deposit (text) on a metallic object

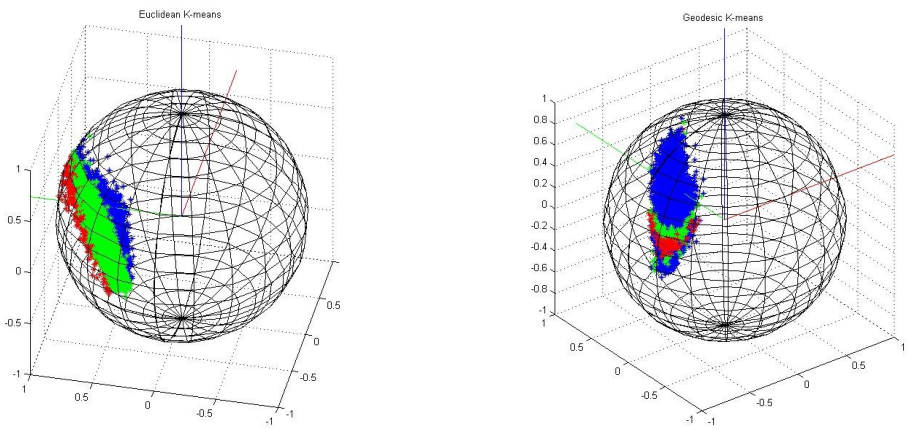

Fig. 7: *Euclidean* and *geodesic* clusters representation in the Poincaré Sphere

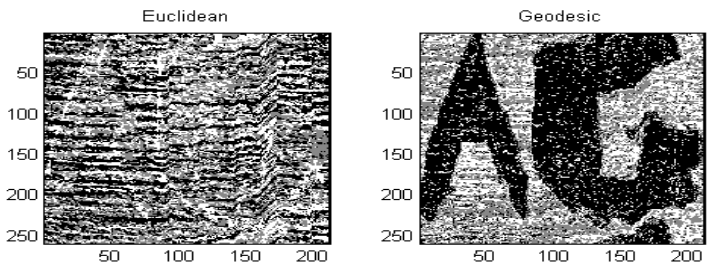

Fig. 8: *Euclidean* and *geodesic* clusters representation in the image using a three grey levels scale

4 – Conclusion.

Working on the assumption that from an experimentalist point of view, pure states of polarization are very unlikely to exist in polarimetric images, we introduced the notion of geodesic distance associated to the polarization space of non-singular coherency matrices. We showed how this distance can be directly related to the Jones calculus. We investigate the question of why use other distances than Euclidean one and what are the appropriate data to analyze with these distances. The question of generalized Gaussian distribution of coherency matrices is also addressed. Eventually, applications on simulated data and real data of a geodesic version of the classical K-means clustering algorithm are presented. The results demonstrate the advantages of the geodesic approach specially when *DOP* is high since for low *DOP*, the non-linear effects are less sensitive.

Appendix A

The mapping φ defined as :

$$\forall g \in GL(2,\mathbb{C}),\quad \forall x \in HDP(2),\quad \varphi(g, x) = g\, x\, g^{\dagger} \tag{A1}$$

is a smooth action group of $GL(2,\mathbb{C})$ on $HPD(2)$. We first demonstrate φ is a mapping of $GL(2,\mathbb{C})$ on $HPD(2)$. If $x \in HPD(2)$, we have $x^{\dagger} = x$ and then:

$$\forall g \in GL(2,\mathbb{C}),\quad \forall x \in HDP(2),\quad \left(g\, x\, g^{\dagger}\right)^{\dagger} = \left(g^{\dagger}\right)^{\dagger} x^{\dagger}\, g^{\dagger} = g\, x\, g^{\dagger} \tag{A2}$$

So, φ(g,x) is Hermitian.

If we consider $v \neq 0 \in \mathbb{C}^2$, $v^{\dagger}(g\, x g^{\dagger})v = (v^{\dagger} g)\, x\, (v^{\dagger} g)^{\dagger} = u^{\dagger} x\, u$ with $u = g^{\dagger} v$. We have $u \neq 0$ since $v \neq 0$ and det(g)$\neq 0$. But for $x \in HPD(2)$ and $\forall\ u \neq 0 \in \mathbb{C}^2$, $u^{\dagger} x\, u > 0$ is the definition of positiveness of x and then φ(g,x) is positive definite. To conclude, φ must satisfy the both conditions of Eqs. (7). The first condition of Eq. (7) is obvious. The proof of the second condition is given by:

$$\forall\left(g_1,g_2\right)\in GL(2,\mathbb{C})\times GL(2,\mathbb{C}),\quad \forall x\in HDP(2),$$
$$\varphi\left(g_1,\varphi\left(g_2,x\right)\right)=\varphi\left(g_1,g_2xg_2^{\dagger}\right)=g_1g_2\,x\,g_2^{\dagger}g_1^{\dagger}=(g_1g_2)\,x\,(g_1g_2)^{\dagger}=\varphi\left(g_1g_2,x\right) \tag{A3}$$

Appendix B

Let G a group acting on a set M , the orbit of $x \in M$ is defined as the set $Gx = \{\varphi(g , x) / g \in G\}$. If $Gx = M$ for some x, then G is said to be *transitive* on M or the G-action to be *transitive* on M. In this case $Gx = M$ for all x. $GL(2,\mathbb{C})$ is transitive on $HPD(2)$, since all Hermitian definite positive matrix $\mathbf{\Phi}$ can be decomposed as:

$$\mathbf{\Phi} = \mathbf{g}\mathbf{g}^{\dagger} = \varphi(g,e) \tag{B1}$$

where e= **Id** is the Identity matrix of rank 2 and $\boldsymbol{g} \in GL(2,\mathbb{C})$.

The *isotropy subgroup* of $x \in M$, is the subgroup of G defined by the set of G -elements that leaves the point x fixed. This subgroup is $G.x = \{ g \in G / \varphi(g , x) = x \}$. For $HPD(2)$ under the mapping φ, the *isotropy subgroup* of **Id** is $U(2)$, the group of unitary matrices of dimension 2, since $\varphi(g , e) = e$ gives from a matrix point of view $\mathbf{g}\,\mathbf{Id}\,\mathbf{g}^{\dagger} = \mathbf{g}\,\mathbf{g}^{\dagger} = \mathbf{Id}$, that is a definition of **g** as a unitary matrix.

The Lie algebra $\mathfrak{g}$ of a Lie group G is the tangent space at the identity with the Lie bracket operation defined by $[\mathbf{X},\mathbf{Y}] = \mathbf{XY}-\mathbf{YX}$ (For more information on this subject, the reader is referred to [20]). Then, it is possible to identify $T_e(M)$. The Lie algebra of matrix Lie group G is the set of all matrices **X** such that exp(t**X**) is in G for all real numbers t (where *exp* stands for the classical exponential of matrix).

For $\mathbf{X} \in H(2)$, the group of Hermitian matrices, $\mathbf{y} = \exp(t\mathbf{X})$ is Hermitian definite positive ($\mathbf{y} \in HPD(2)$). This property is straightforward since:

$$\exp(t\mathbf{X}) = \exp[\,t\,(\mathbf{X}/2 + \mathbf{X}/2)] = \exp[\,t\,(\mathbf{X}^{\dagger}/2 + \mathbf{X}/2)] = \exp(t\,\mathbf{X}/2)^{\dagger}\exp(t\,\mathbf{X}/2)$$

(since the classical relation exp(**A**+**B**)=exp(**A**)exp(**B**) if [**A**,**B**]=0 is verified). This last expression is obviously an hermitian definite positive matrix. Then, $H(2)$ the manifold of Hermitian matrices is the tangent space of $HPD(2)$ at **Id**.

For the smooth action group of a Lie group G on a manifold M defined by the mapping $\varphi(g , x)$, the differential $d\varphi(g , x)$ of φ at a point $p \in M$ will be a map of $T_p(M)$ to $T_{\varphi(g,p)}(M)$. Thus, the mapping φ associated with Eq. (3), gives $d\varphi(g , X) = \mathbf{g}\,\mathbf{X}\,g^{\dagger}$, that is a tangent vector at $gpg^{\dagger}$ to $HPD(2)$. It is straightforward to check that all the resulting tangent spaces are $H(2)$ since each $\mathbf{p} \in HPD(2)$ can be write as $\mathbf{p} = \mathbf{g}\,\mathbf{g}^{\dagger}$ then, it is always possible to define a map $d\varphi(g^{-1}, X) = \mathbf{g}^{-1}\,\mathbf{X}\,\mathbf{g}^{-\dagger}$ that is a tangent vector at **Id** to $HPD(2)$. According to Theorem 1, the corresponding transformations are isometries and we have:

$$\left\langle X_p,Y_p\right\rangle_p = \left\langle d\varphi\left(g^{-1},X_p\right),d\varphi\left(g^{-1},Y_p\right)\right\rangle_{Id} \tag{B2}$$

For the geodesic with $\mathbf{Y} = d\varphi(g^{-1}, X_p) = \mathbf{g}^{-1}\,\mathbf{X_p}\,\mathbf{g}^{-\dagger}$

$$\varphi\left[g^{-1},\gamma_{(p,X)}(t)\right] = \gamma_{(Id,Y)}(t) = \exp(t\mathbf{Y}) \tag{B3}$$

then

$$\gamma_{(p,X)}(t) = \varphi\left[g,\gamma_{(Id,Y)}(t)\right] = \mathbf{g}\exp(t\mathbf{Y})\,\mathbf{g}^{\dagger} \tag{B4}$$

As $\mathrm{Exp}(X_p) = \gamma_{(p,X)}(1) = x$, we have:

$$\gamma_{(p,X)}(1) = \mathbf{g}\exp(\mathbf{Y})\,\mathbf{g}^{\dagger} \tag{B5}$$

Appendix C

Algorithm for the generation of coherency matrices with a generalized normal distribution of small variance.

Input : $\mathbf{\Phi}$ the mean coherency matrix of the distribution and $\mathbf{\Lambda}$ the covariance matrix
output : $\mathbf{\Phi_k}$
1 – Perform the Cholesky decomposition $\mathbf{\Lambda} = \mathbf{HH^T}$
2 – Generate a random vector $\mathbf{z} \in \mathbb{R}^4$ with independent and identically distributed components and with zero mean.
3 – Reshape $\mathbf{X_k} = \mathbf{Hz}$ as an Hermitian matrix (an element of the tangent space of *HPD(2),* see Appendix B)
4 – Compute $\mathbf{\Phi_k} = \left[\exp\left(-\mathbf{\Phi}^{-1}\mathbf{X_k}\right)\mathbf{\Phi}\right]^{-1}$

References.

1 J. Scott Tyo, Dennis L. Goldstein, David B. Chenault, and Joseph A. Shaw, "Review of passive imaging polarimetry for remote sensing applications", Appl. Opt. **45**, 5453-5469 (2006)
2 Vincent Nourrit, Juan M. Bueno, Brian Vohnsen, and Pablo Arta, "Nonlinear registration for scanned retinal images: application to ocular polarimetry", Appl. Opt. **47**, 5341-5347 (2008)
3 A. N. Yaroslavsky, V. Neel, and R. R. Anderson, "Fluorescence polarization imaging for delineating nonmelanoma skin cancers", Opt. Lett. **29**, 2010-2012 (2004)
4 F Boulvert, B Boulbry, G Le Brun, B Le Jeune, S Rivet and J Cariou, "Analysis of the depolarizing properties of irradiated pig skin", J. Opt. A: Pure Appl. Opt. **7**, 21–28 (2005)
5 Steve Guyot, Makrina Anastasiadou, Eric Deléchelle, and Antonello De Martino, "Registration scheme suitable to Mueller matrix imaging for biomedical applications", Opt. Express **15**, 7393-7400 (2007)
6 Axel Kreuter, Matthias Zangerl, Michael Schwarzmann, and Mario Blumthaler, "All-sky imaging: a simple, versatile system for atmospheric research", Appl. Opt. **48**, 1091-1097 (2009)
7 Patrick Terrier, Vincent Devlaminck, and Jean Michel Charbois, "Segmentation of rough surfaces using a polarization imaging system", J. Opt. Soc. Am. A **25**, 423-430 (2008)
8 Olivier Morel, Christophe Stolz, Fabrice Meriaudeau, and Patrick Gorria, "Active lighting applied to three-dimensional reconstruction of specular metallic surfaces by polarization imaging", Appl. Opt. **45**, 4062-4068 (2006)
9 Arnaud Bénière, François Goudail, Mehdi Alouini, and Daniel Dolf, "Degree of polarization estimation in the presence of nonuniform illumination and additive Gaussian noise", J. Opt. Soc. Am. A **25**, 919-929 (2008)
10 Jawad Elsayed Ahmad and Yoshitate Takakura, "Error analysis for rotating active Stokes-Mueller imaging polarimeters", Opt. Lett. **31**, 2858-2860 (2006)
11 Christian Brosseau, *Fondamentals of polarized light: a statistical optics approach*, (John Wiley, New York, 1998)
12 R. C. Jones, "A new calculus for the treatment of optical systems", J. Opt. Soc. Am. **31**, 488-493 (1947)
13 E. Wollf, "Coherence properties of partially polarized electromagnetic radiation", Nuevo Cimento **13**, 1165-1181 (1959)
14 K. Kim, L. Mandel, and E. Wolf, "Relationship between Jones and Mueller matrices for random media", J. Opt. Soc. Am. A **4**, 433–437 (1987)
15 R. Simon, "The connection between Mueller and Jones matrices of polarization optics", Opt. Com. **42**, 293-297 (1982)
16 E.S. Fry, G.W. Kattawar, "Relationships between elements of the Stokes matrix", Appl. Opt. **20**, 2811-2814 (1981)
17 S.R. Cloude, "Group theory and polarisation algebra", Optik **75**, 26-36 (1986)
18 Williams M. Boothby, *An introduction to Differentiable Manifolds and Riemannian Geometry-Second Edition*, (Academic Press, Inc., Orlando, 1986)
19 P.T. Fletcher and S.C. Joshi, "Principal geodesic analysis on symmetric spaces: statistics of diffusion tensors", *in Proc. of CVAMIA and MMBIA Workshops*, Springer, Prague, Czech Republic, (2004), pp.87-98
20 Sigudur Helgason, *Differential Geometry and Symmetric spaces*, (Academic Press, New York, 1962)

21 Ragini Verma, Parmeshwar Khurd and Christos Davatzikos,“On Analyzing Diffusion Tensor Images by Identifying , Manifold Structure Using Isomaps”,IEEE Trans. Med. Img. **26**, 772-778 (2007)

22 J. B. Tenenbaum, V. de Silva, and J. C. Langford, “A global geometric framework for nonlinear dimensionality reduction”, Science **290**, 2319–2323 (2000)

23 C. J. C. Burges, “Geometric Methods for feature extraction and dimensional reduction”, in Data Mining and Knowledge Discovery Handbook: A Complete Guide for Practitioners and Researchers, (O.Maimon, Ed. Norwell, MA: Kluwer, 2005).

24 M. Fréchet, "Les éléments aléatoires de nature quelconque dans un espace distancié", Ann. Inst. H. Poincaré **10**, 215-310 (1948)

25 Xavier Pennec, "Intrinsic Statistics on Riemannian Manifolds: Basic tools for Geometric Measurements", J. Math Imaging Vis. **25**, 127-154 (2006)

26 H. Karcher, "Riemannian center of mass and mollifier smoothing", Comm. Pure Appl. Math. **30**, 509-541 (1977)

27 U. Grenander, “Probabilities on Algebraic Structures”, (John Wiley and Sons, 1963).

28 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 Vis **25**: 423–444 (2006)

29 John A. Hartigan, *Clustering algorithms.Wiley Series in Probability and Mathematical Statistics*, (John Wiley and Son, New York, 1975)