 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717

Article

Representation and Characterization of Nonstationary Processes by Dilation Operators and Induced Shape Space Manifolds

Maël Dugast , Guillaume Bouleux and Eric Marcon

Univ Lyon, UJM-Saint Etienne, INSA Lyon, DISP, F-69621, Villeurbanne, France;
{mael.dugast,guillaume.bouleux,eric.marcon}@insa-lyon.fr
* Correspondence: guillaume.bouleux@insa-lyon.fr; Tel.: +33-477448920

Abstract: We proposed in this work the introduction of a new vision of stochastic processes through geometry induced by dilation. The dilation matrices of a given process are obtained by a composition of rotation matrices built in with respect to partial correlation coefficients. Particularly interesting is the fact that the obtention of dilation matrices is regardless of the stationarity of the underlying process. When the process is stationary, only one dilation matrix is obtained and it corresponds therefore to Naimark dilation. When the process is nonstationary, a set of dilation matrices is obtained. They correspond to Kolmogorov decomposition. In this work, the nonstationary class of periodically correlated processes was of interest. The underlying periodicity of correlation coefficients is then transmitted to the set of dilation matrices. Because this set lives on the Lie group of rotation matrices, we can see them as points of a closed curve on the Lie group. Geometrical aspects can then be investigated through the shape of the obtained curves, and to give a complete insight into the space of curves, a metric and the derived geodesic equations are provided. The general results are adapted to the more specific case where the base manifold is the Lie group of rotation matrices, and because the metric in the space of curve naturally extends to the space of shapes, this enables a comparison between curves' shapes and allows then the classification of processes' measure.

Keywords: nonstationary processes; spectral measure; differential geometry; shape manifold; square root velocity function; Lie group

---

1. Introduction

The analysis and/or the representation of nonstationary processes has been tackled for 4 or 5 decades now by time-scale/time-frequency analysis [4,21], by Fourier-like representation when the processes belong to the periodically correlated (PC) subclass [25,40], or by partial correlation coefficients (pacors) series [20,30], to cite a few. One of the advantages of dealing with parcors resides in their strong relation to the measure of the process by the one-to-one relation with correlation coefficients [18,55]. They consequently appear explicitly in the Orthogonal Polynomial on the Real Line/Unit Circle decomposition of the measure [11,47] and are the elements for the construction of dilation matrices that appear in the CMV/GGT [46], for the Schur flows problem with upper Hessenberg matrices [1] that are also seen in the literature as evolution operators[47] or shift operator [35], and finally appear in the state-space representation [15,17]. The dilation theory takes its roots from the operator theory [51], which bridges the process' measure and unitary operators. In its simplest version, the dilation theory corresponds to Naimark dilation [3,51], and states that given a sequence of correlation coefficients, there exists a unitary matrix W such that $R_n \triangleq (1\,0\,0\cdots)W^n(1\,0\,0\cdots)^T$ where $\cdot^T$ denotes the transposition. When the process is not stationary, its associated correlation matrix is no more Toeplitz structured, a set of matrices is required [15] and the previous expression becomes $R_{i,j} \triangleq (1\,0\,0\cdots)W_{i+1}W_{i+2}\cdots W_j(1\,0\,0\cdots)^T$. The matrices W_i are theoretically understood as infinite rotation matrices, which become finite when the correlation coefficients sequence is itself

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

finite. In that particular case, the matrices W_i belong to $SO(n)$ or $SU(n)$, the special orthogonal or unitary group, respectively, and the process' measure is totally described by the set of W_i. As a consequence, the measure of the process is beautifully characterised for the nonstationary case, by a sampled trajectory induced by the dilation matrices on the appropriate Lie group. When the process is periodically correlated, the sequence of parcors inherits the periodicity and the sequence of dilation matrices becomes periodic as well, we consequently obtain a closed path as illustrated in Figure1. Characterising the time-varying measure of the process is now tackled by studying curves (or sampled curves) on special groups.

Information geometry is now a fundamental approach to describing stochastic processes [34]. The second-order statistical properties/moments may be analysed, characterised and compared [5,8] to improve the estimation [39,50] or classification of different processes [28]. When dealing with density estimation [26], the space of $n \times n$ symmetric matrices $\mathcal{S}ym(n)$ is generally preferred, and many developments have been proposed under the semi-positive-definite (SPD) assumption [14,41,42,48] for which the set of SPD matrices constitute a convex half-cone in the vector space of matrices. This leads to giving more insights into the Fisher information metric [14,26] or the Wasserstein metric [29] and coping with optimal mass transportation problems [7]. Many efforts have also been made in the last decade to exploit the hyperbolic geometry structure not of the correlation matrices directly but of the related parcors when obtained in stationary conditions [2,6,16,19,55]. As the Kullback–Leibler divergence let do, the comparison of stationary processes is then made by comparing curves, whose sampled points are parcors sequences, defined on several copies of the Poincaré disk through geodesics deformation. Treating the nonstationary case has not been tackled to our knowledge with the previous mentioned approaches. In this paper, we hope to initiate interest in filling this gap by extending the representation and the characterisation of processes' measure in nonstationary context, first in using the dilation theory approach to give sampled points and then in giving the prescribed geodesics equations used for curve or path comparisons in the Lie group.

To support the reader, some insights on dilation theory are given in Section 2. Practical implementations of dilation matrices according to the operator theory approach [3,15] or the lattice filter structure approach [27,44] are also discussed and the strong connection between parcors and the dilation matrices is emphasised. Section 3 focuses on the geometry of the curves induced by the dilation on particular manifolds. The general framework is first introduced by recalling concepts of distances and shape of curves when the ambient space is not flat. Next, the square root velocity (SRV) functions are developed and adapted to the Lie group, and a procedure to compare nonstationary processes through their time evolution trajectory is presented. Finally, a conclusion is drawn in Section 4 and the reader will find some technical tools in the Appendix section.

2. The structure of semi-positive-definite matrices and the dilation theory

2.1. The theory of dilation and the interaction with

Let us give some insights into the dilation theory. In its fundamental definition, the dilation theory consists of a Hilbert space $\mathcal{H}$ and an operator-valued function f, *i.e.* an $\mathcal{L}(\mathcal{H})$-valued function, to find a larger Hilbert space H and an other application $\mathcal{F}$ such that f is the orthogonal projection of $\mathcal{F}$:

$$f(t) = P_{\mathcal{H}}\mathcal{F}(t), \qquad t \in \mathbb{Z} \tag{1}$$

where $P_{\mathcal{H}}$ denotes the orthogonal projection onto the Hilbert space $\mathcal{H}$. The ideas of the dilation theory are :

- there exists a larger space from which the original function (or matrix) is deduced
- we can choose the "dilated" function to be simpler. For instance, when dealing with matrices, each of its coefficients can be expressed as the projection of a larger unitary matrix. In this case,

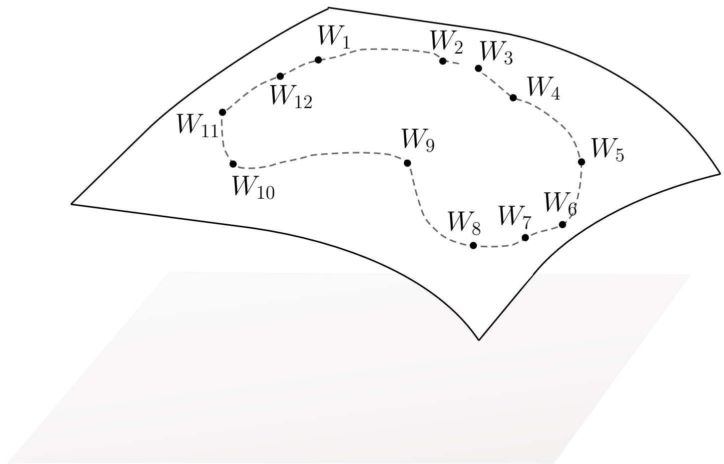


Figure 1. Illustration of a sampled closed trajectory drawn in $SO(n)$ or $SU(n)$ that materialises the time varying of the PC measure for a stochastic process. Each W_i is a dilation matrix built through the parcors.

we obtain a unitary dilation. This approach has been for example developed in [33,36] and [37] for the stationary dilation of periodically-correlated processes.

2.1.1. Dilation and rotation of contractions

For an operator T on a Hilbert space $\mathcal{H}$, we denote by T^* the adjoint operator, *i.e.* the operator on $\mathcal{H}$ such that $\langle Tx, y\rangle = \langle x, T^*y\rangle$ for all $x, y \in \mathcal{H}$. An operator $T \in \mathcal{L}(\mathcal{H})$ is said to be a contraction if $|| \; T \; || \leq 1$ where $|| \cdot ||$ is the operator norm. We deduce the expression for the defect operator $D_T = (I - T^*T)^{1/2}$ and its adjoint $D_{T^*} = (I - TT^*)^{1/2}$.
One of the easiest results is that, given a contraction Γ, the aforementioned unitary Julia operator

$$J(\Gamma) = \begin{pmatrix} \Gamma & D_{\Gamma^*} \\ D_\Gamma & -\Gamma^* \end{pmatrix} \tag{2}$$

satisfies, for all $n \in \mathbb{N}$

$$\Gamma^n = \begin{pmatrix} 1 & 0 \end{pmatrix} J(\Gamma)^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{3}$$

In other words, the elementary rotation of a contraction, called consequently the Julia operator, also corresponds to the unitary dilation operator of the contraction. Note that the Julia operator is sometimes called the Halmos extension [35] of a contraction.

2.1.2. Dilation and isometries

Following the idea and the formulation of Naimark, the dilation theory can be restated in terms of dilation of the sequence of operators or sequence of numbers when the dimension of the underlying Hilbert space is 1. Recall that a sequence of operators $\{R_n\}_{n=1}^{\infty}$ acting on $\mathcal{H}$ is said to be positive if

$$\sum_{i,j=0}^{+\infty} \langle R_{i-j} h_i, h_j \rangle \geq 0 \quad for\ all\ h_i \in \mathcal{H}_i. \tag{4}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

Assuming now that $R_n^* = R_{-n}$ and $R_0 = I$, leads to the following Toeplitz matrix:

$$R^{(m)} = \begin{pmatrix} I & R_1 & \cdots & R_{m-1} \\ R_1^* & I & \cdots & R_{m-2} \\ . & . & \cdots & . \\ . & . & \cdots & . \\ R_{m-1}^* & R_{m-2}^* & \cdots & I \end{pmatrix} \tag{5}$$

which is positive-definite. Remark that this matrix can be seen as the correlation matrix of a stationary process, as it is positive and Toeplitz [3,12,54]. Owing to this property, we obtain the following relation:

$$R_n = P_{\mathcal{H}} U^n \mid_{\mathcal{H}}, \quad \textit{for all } n \geq 0 \text{ and } U \text{ an isometry on } \mathcal{K} \tag{6}$$

as a result of the Naimark dilation theorem. Furthermore, if $\mathcal{K} = \bigvee_{n \geq 0} U^n \mathcal{H}$ then U is unique up to an isomorphism.

2.1.3. Dilation and measure

From Bochner's theorem, we known a matrix of type (5) can be seen as the Fourier coefficient of a given positive Borelian measure. This is also known as the moment or trigonometric problem [15]. Therefore, we can restate the dilation problem in terms of measure. If we denote by E_λ an operator-valued distribution function on $[0, 2\pi[$, then the function

$$R_n = \int_0^{2\pi} e^{in\lambda} dE_\lambda. \tag{7}$$

This function is positive-definite and shows the strong correspondence between the spectral measure and the dilation theory. There hence exists a unitary operator on a Hilbert space $\mathcal{K}$ such that $R_n = P_{\mathcal{H}} U(n)$ where $P_{\mathcal{H}}$ stands for the orthogonal projection. With the spectral representation of unitary operators, $U = \int_0^{2\pi} e^{i\lambda} dE_\lambda$ and we have

$$\int_0^{2\pi} e^{in\lambda} d\langle E_\lambda u, v\rangle = \int_0^{2\pi} e^{in\lambda} d\langle F_\lambda u, v\rangle \tag{8}$$

or, in an equivalent form :

$$E_\lambda = P_{\mathcal{H}} F_\lambda. \tag{9}$$

Note that the operator-valued measure F_λ is in fact an orthogonal projection-valued measure because all its increments are orthogonal.
With dilation matrices having been introduced, we give now in the next section a methodology to understand how they are obtained.

2.2. Construction of Dilation Matrices

As mentioned previously, given an SPD matrix $R = (R_{i,j})_{i,j \in \mathbb{N}}$, it is possible to build a sequence of matrices $\{W_i\}_{i \in \mathbb{N}}$ such that $R_{i,j} = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix} W_i W_{i+1} \cdots W_{j-1} \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \end{pmatrix}^T$ by a two-step procedure. For the first step, the following theorem is needed [15] :

Theorem 1 (Structure of a positive-definite block matrix)**.** *Let X and Z be positive operators in $\mathcal{L}(\mathcal{H}_X)$ and $\mathcal{L}(\mathcal{H}_Z)$ respectively. Then the following are equivalent :*

- *The operator $A = \begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix}$ is positive*

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

- *There exists a unique contraction* Γ *in* $\mathcal{L}(\mathcal{R}(Z), \mathcal{R}(X))$ *such that*

$$Y = X^{1/2}\Gamma Z^{1/2} \tag{10}$$

Proof. Annexe A □

Let us now apply this relation repeatedly on an SPD matrix. To fix ideas, let the 3×3 (block-)matrix be

$$R = \begin{pmatrix} R_{1,1} & R_{1,2} & R_{1,3} \\ R_{1,2}^* & R_{2,2} & R_{2,3} \\ R_{1,3}^* & R_{2,3}^* & R_{3,3} \end{pmatrix} \tag{11}$$

and apply Theorem 1 to $\begin{pmatrix} R_{1,1} & R_{1,2} \\ R_{1,2}^* & R_{2,2} \end{pmatrix}$, $\begin{pmatrix} R_{2,2} & R_{2,3} \\ R_{2,3}^* & R_{3,3} \end{pmatrix}$ and finally to $\begin{pmatrix} R_{1,2} & R_{1,3} \end{pmatrix}$. Note that when a square root of a (block-)matrix has to be chosen, it is done according to the Schur decomposition given in Annexe A. At each step, a contraction $\Gamma_{i,j}$ is generated with respect to the indices of the upper and lower (block-)matrices of the main diagonal, *e.g.* $\Gamma_{1,2}$ for the first $\begin{pmatrix} R_{1,1} & R_{1,2} \\ R_{1,2}^* & R_{2,2} \end{pmatrix}$ (block-)matrix. We thus obtain a one-to-one correspondence between the SPD matrix R and the set of contractions $\{\Gamma_{i,j}\}_{i=1,2\ j=3}$. Regarding the huge work of Constantinescu [15], we will called these contractions the Schur-Constantinescu parameters. Considering now unit variance and arbitrary size $n \times n$ for the SPD matrix, allows us to write the correspondence as:

$$\begin{pmatrix} I & R_{1,2} & & R_{1,n} \\ R_{1,2}^* & I & \ddots & \\ & \ddots & \ddots & R_{n-1,n} \\ R_{1,n}^* & & R_{n-1,1}^* & I \end{pmatrix} \longleftrightarrow \begin{pmatrix} 0 & \Gamma_{1,2} & \Gamma_{1,3} & & \cdots & \Gamma_{1,n} \\ 0 & 0 & \Gamma_{2,3} & \Gamma_{2,4} & \cdots & \Gamma_{2,n} \\ \vdots & & \ddots & \ddots & \ddots & \\ & & & & & \Gamma_{n-2,n} \\ & & & & 0 & \Gamma_{n-1,n} \\ 0 & 0 & \cdots & & & 0 \end{pmatrix}. \tag{12}$$

Once (12) is established, each dilation matrix W_i is built-up as a product of Givens rotations of a sequence of Schur-Constantinescu parameters in the following way:

$$W_i = G(\Gamma_{i,i+1})G(\Gamma_{i,i+2}) \cdots G(\Gamma_{i,j}), \tag{13}$$

where $G_{\Gamma_{i,i+l}}$ denotes the Givens rotation of $\Gamma_{i,i+l}$ as follows:

$$G(\Gamma_{i,i+l}) = I \oplus \begin{pmatrix} \Gamma_{i,i+l} & D_{\Gamma_{i,i+l}^*} \\ D_{\Gamma_{i,i+l}} & -\Gamma_{i,i+l}^* \end{pmatrix} \oplus I. \tag{14}$$

When the SPD matrix is Toeplitz, which correspond to a stationary underlying process, then all dilation matrices W_i are identical and they take the form

$$W_i = U = \begin{pmatrix} \Gamma_1 & D_{\Gamma_1^*}\Gamma_2 & D_{\Gamma_1^*}D_{\Gamma_2^*}\Gamma_3 & D_{\Gamma_1^*}D_{\Gamma_2^*}D_{\Gamma_3^*}\Gamma_4 & \cdots & \\ D_{\Gamma_1} & -\Gamma_1^*\Gamma_2 & -\Gamma_1^*D_{\Gamma_2^*}\Gamma_3 & -\Gamma_1^*D_{\Gamma_2^*}D_{\Gamma_3^*}\Gamma_3 & \cdots & \\ 0 & D_{\Gamma_2} & -\Gamma_2^*\Gamma_3 & -\Gamma_2^*D_{\Gamma_3^*}\Gamma_4 & \cdots & \\ 0 & 0 & D_{\Gamma_3} & -\Gamma_3^*\Gamma_4 & \cdots & \\ 0 & 0 & 0 & D_{\Gamma_4} & \cdots & \\ \cdot & \cdot & \cdot & \cdot & & \cdots \\ \cdot & \cdot & \cdot & \cdot & & \cdots \end{pmatrix} \tag{15}$$

 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

which is nothing less than the Naimark dilation introduced in the first part, *i.e.* $R_{i,j} = R_{j-1} = [1\ 0\ 0\ \cdots]U^{j-i}[1\ 0\ 0\ \cdots]^T$. For the sake of completeness, we give the correspondence between the coefficients of the SPD matrix (the left-hand side of (12)) and the Schur-Constantinescu parameters:

Theorem 2. *The matrix $R^{(n)} = [R_{k,j}]^n_{k,j=1}$, satisfying $R^*_{j,k} = R_{k,j}$ is positive if and only if*

- *$R_{kk} \geqslant 0$ for all k*
- *there exists a family $\{\Gamma_{k,j} \mid k, j = 1, \cdots n, k \leqslant j\}$ of contraction such that*

$$R_{k,j} = B^*_{k,k}(L_{k,j-1}U_{k+1,j-1}C_{k+1,j} + D_{\Gamma^*_{k,k+l}} \cdots D_{\Gamma^*_{k,j-l}} \Gamma_{k,j} D_{\Gamma_{k+1,j}} \cdots D_{\Gamma_{j-1,j}})B_{j,j} \tag{16}$$

where $B_{k,k}$ is any square root of $R_{k,k}$

and

$$L_{k,j} = [\Gamma_{k,k+1} \quad D_{\Gamma^*_{k,k+l}}\Gamma_{k,k+2} \quad \cdots \quad D_{\Gamma^*_{k,k+l}} \cdots D_{\Gamma^*_{k,j-1}}\Gamma_{k,j}] \tag{17}$$

a row contraction associated to the set of parameters $\{\Gamma_{k,m} \mid k < m \leq j\}$,

$$C_{k,j} = [\Gamma_{j-1,j} \quad \Gamma_{j-2,j}D_{\Gamma_{j-1,j}} \quad \cdots \quad \Gamma_{k,j}D_{\Gamma_{k+1,j}} \cdots D_{\Gamma_{j-1,j}}]^T \tag{18}$$

a column contraction associated to the set of parameters $\{\Gamma_{m,j} \mid m = j-1, \cdots k\}$, and finally

$$U_{k,j} = G(\Gamma_{k,k+1})G(\Gamma_{k,k+2}) \cdots G(\Gamma_{k,k+j}) \left(U_{k+1,j} \oplus I\right) \tag{19}$$

Proof. This theorem is proved in [15]. □

A different approach leading to the same results can be found in [52], using directly the Kolmogorov decomposition. In [27] the Naimark dilation is constructed using the lattice filter and finally applications of this decomposition in quantum mechanics are to be found in [53,54] for example.

3. Analysis of curves on a manifold induced by the dilation

Parcors, composing dilation matrices, have already been given a geometrical point of view, as, for example, in [55] where the sequence of parcors associated with a stationary process is seen as a point onto the Poincaré polydisk $\mathcal{P}^n$, that is, the product of the Poincaré disk. To give geometrical settings, a distance to characterise individual parcors is then proposed and discussed. In [31], a stochastic process is studied under the local stationarity assumption. To each stationary slice of the process corresponds a sequence of parcors, represented as a point in the Poincaré polydisk $\mathcal{P}^n$ as well. A trajectory is then generated on that space which materialises a curve on the manifold $\mathcal{P}^n$. The underlying computations are quite intricate because of the product manifolds, and the question of nonstationarity arises. Based on the works of Le Brigant [31,32], Celledoni *et al.* [13] and Zhang *et al.* [57], we propose then to give a particular attention to this question. We first make use of the dilation theory introduced in Section 2. When the process under study is nonstationary, a set of matrices W_i is obtained. The basic idea for having geometric information on the nonstationary process is therefore to characterize the trajectory formed by the set of dilation matrices. These matrices are theoretically operators of infinite dimension, but as we dispose of only a finite set of parcors, the theoretical matrices of (15) are truncated. Matrices respecting (15) are general rotation matrices that become perfect rotation operators belonging to $SO(n)$ for real processes and $SU(n)$ when dealing with complex processes, when their dimensions are reduced to $n \times n$. Our aim is finally to analyse those curves living on the Lie group of rotation matrices and emphasise the geometry or, more precisely, the intrinsic geometry formulation of these objects. For example, we aim at comparing different curves coming from different processes or at resuming many realisations of a stochastic process (multiple measurements) through the computation of the mean of the associated several curves. The question as to computation complexity still exists,

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

but many results have been proposed recently to overcome this difficulty and to propose closed-form formulations. In particular, it is predicated to extract the shape of the trajectory for it contains the essentials, in topologic sense, information.

To allow the curves comparison, we have based our development on the works of Le Brigant [31] and Celledoni *et al.* [13]. First, we define the manifold $\mathcal{M}$ given by the set of all curves in the base manifold. This leads to another space, the shape space, for which the manifold $\mathcal{M}$ will be a fiber bundle. We dispose then of a metric in $\mathcal{M}$ from which a metric on the shape space is deduced. These steps are now explained in the following.

3.1. Basic Outline of Geometry

Curves of interest are those living in the Lie group of real rotation matrices; this yields $c : [0,1] \to SO(n)$. For the sake of clarity, we suppose that c is continuous, we will come back to the case of discrete curves later. To study the geometrical features of such curves, we interest ourselves with the set of all curves lying in $SO(n)$ (where $SO(n)$ is seen as a manifold) with nonvanishing velocity, *i.e.* $\mathcal{M} = \{c \in \mathcal{C}^\infty([0,1], SO(n)) : c'(t) \neq 0\ \forall t\}$, this is in fact a sub-manifold of $\mathcal{C}^\infty([0,1], SO(n))$. A curve c is thus a particular point in $\mathcal{M}$. The tangent space at a curve c is given by

$$T_c\mathcal{M} = \left\{ v \in \mathcal{C}^\infty([0,1], TSO(n)) : v(t) \in T_{c(t)}SO(n) \right\} \tag{20}$$

where $TSO(n)$ denotes the tangent bundle of the base manifold $SO(n)$. Note that a tangent vector is a curve in the tangent space of $SO(n)$. In this manifold, the expression of distances and, thus, geodesics depends on the chosen metric. When comparing two curves, it is natural that the distance between these two curves should remain the same if the curves are only reparametrised, that is, if we define other curves that pass through the same points than the original curves but at different speeds. When the curve is discretised as we will see in the sequel, doing a reparametrisation is equivalent to changing the chosen points (see Figure2). A reparametrisation is represented by increasing diffeomorphism $\phi \in \mathcal{D} : [0,1] \to [0,1]$ acting on the right of the curve by composition. In other words, we required that the Riemannian metric g on $\mathcal{M}$ satisfies the following property:

$$g_{c\circ\phi}(u \circ \phi, v \circ \phi) = g_c(u, v) \tag{21}$$

for all $c \in \mathcal{M}$, $u, v \in T_c\mathcal{M}$ and $\phi \in \mathcal{D}$. This property is called reparametrisation invariance. We insist

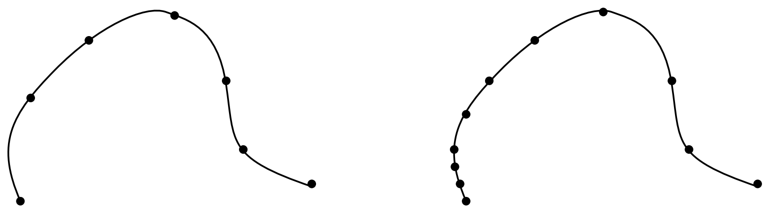


Figure 2. Example of a reparametrisation of a curve. Here, it consists in changing the discretisation with nonlinear time sample.

on the fact that g is the metric on $\mathcal{M}$, the space of all curves on $SO(n)$ and not on $SO(n)$ itself. In terms of distances, this gives

$$d_{\mathcal{M}}(c_0 \circ \phi, c_1 \circ \phi) = d_{\mathcal{M}}(c_0, c_1) \tag{22}$$

where $d_{\mathcal{M}}$ denote the distance on $\mathcal{M}$ corresponding to the metric g. The reparametrisation introduced above induces an equivalence relation between points in $\mathcal{M}$ such that

$$c_0 \sim c_1 \iff \exists \phi \in \mathcal{D} : c_0 = c_1 \circ \phi. \tag{23}$$

With this equivalence relation, a quotient space can be constructed as the collection of equivalence classes, it is named the shape space and has the following writing:

$$\mathcal{S} = \mathcal{M}/ \sim, \text{ or } \mathcal{S} = \mathcal{M}/\mathcal{D}. \tag{24}$$

A distance function on the shape space is obtained from the distance on $\mathcal{M}$ as follows:

$$d_{\mathcal{S}}([c_0],[c_1]) = \inf_{\phi \in \mathcal{D}} d_{\mathcal{M}}(c_0, c_1 \circ \phi) \tag{25}$$

where $[c_0]$ and $[c_1]$ are representatives of the equivalence classes of c_0 and c_1 respectively. It can be shown that this distance is independent of the choice of the representatives. It is in fact inherited from the fiber bundle structure $\pi = \mathcal{M} \to \mathcal{S}$. As closed curves are of main interest in this work, we can also define the set

$$\mathcal{M}^c = \left\{ c \in \mathcal{C}([0,1], SO(n)) : c'(t) \neq 0, c(0) = c(1) \right\}. \tag{26}$$

Basically, the closure of a curve just imposes the equality of the first and the last point of it, and not of their first derivative. Consequently, $\mathcal{M}^C$ turns into

$$\mathcal{M}^{c+} = \left\{ c \in \mathcal{C}([0,1], SO(n)) : c'(t) \neq 0, c(0) = c(1), c'(0) = c'(1) \right\}. \tag{27}$$

We need now to introduce the Square Root Velocity function (SRV function) [49], in which a curve is represented by its starting point and its normalised velocity at each time t. There are several possibilities to define the SRV of a curve. The more general definition is the following

$$\begin{aligned} F : \mathcal{M} &\to SO(n) \times T\mathcal{M} \\ c &\to \left(c(0), q = \frac{c'}{\sqrt{\| c' \|}} \right). \end{aligned} \tag{28}$$

However we can go further and benefit from the specific case of Lie group. In this section, we will denote the base manifold $G = SO(n)$ to emphasise its group structure, and g an element of the group. As in [13], we consider only curves that start at the identity, this is because other curves can be reduced to this case by right or left translation. In these settings, it is interesting to turn the SRV function into the Transported SRV function (TSRV). This is basically the SRV that has been parallel transported to a reference point. Different versions have been given in [9], [13] or [57] which differ in the choice of their reference point. For our case of study, the identity is our natural curve starting point and is thus a particularly good choice for being the reference point. In a Lie group, a parallel transport operation can be defined, here again, by the right (or left) translation. This justifies that we can take, as suggested in [13] a TSRV function of the following form:

$$\begin{aligned} F_{Lie} : \mathcal{C}^\infty([0,1], G) &\longrightarrow SO(n) \times \{ q \in \mathcal{C}^\infty([0,1], \mathfrak{g}), q(t) \neq 0, \ \forall t \in [0,1] \} \\ F_{Lie}(c)(t) = (c(0), q(t)) &= \left(c(0), \frac{R^{-1}_{c(t)*}(c'(t))}{\sqrt{\| c'(t) \|}} \right) = \left(c(0), \frac{T_c^{c(t) \to I}(c'(t))}{\sqrt{\| c'(t) \|}} \right), \end{aligned} \tag{29}$$

where $\mathfrak{g}$ is the Lie algebra, R is the right translation on the group, $R_{g_1}(g_2) = g_2 g_1$, $R_{g*} = T_e R_g$ is the tangent map at the identity, $\| \cdot \|$ is a norm induced by a right-invariant metric on G, and $T_c^{c(t) \to I}$ denotes the parallel transport from $c(t)$ to the identity according to the curve c. A curve is now represented as an element of the tangent bundle $(c(0), q(t)) \in M \times TM$ (recall that q draws a curve in the tangent bundle), and $c(0)$ is the identity element of the Lie group. The inverse of the SRV function is then straightforward: for every $q \in \mathcal{C}^\infty([0,1], T\mathcal{M})$, there exists a unique curve c such that $F(c_i) = q_i$ and $c(t) = \int_0^t q(r) \| q(r) \| \, dr$ where $\| \cdot \|$ is the norm in $SO(n)$.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

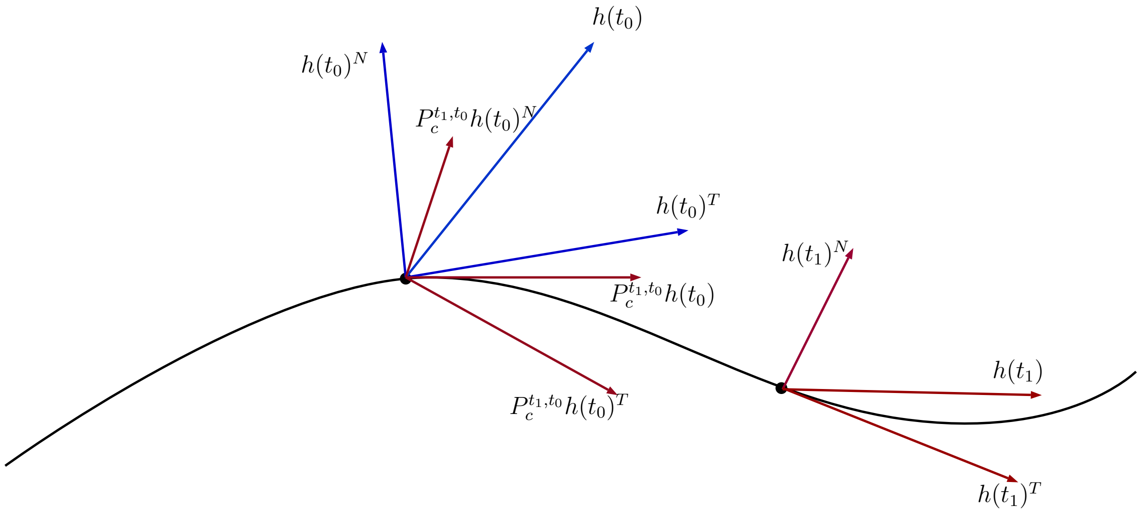


Figure 3. The inner product measures the angle between a frame at a given point and the parallel transport version of this frame at a latter time.

149 *3.2. Metric and distance over $\mathcal{M}$ and $\mathcal{S}$*

We now give insights on a relevant metric that should be used on $\mathcal{M}$ to compare different closed trajectories. The idea is to have a metric on $\mathcal{M}$ that induced a "coherent" distance on the shape space $\mathcal{S}$. The following development and expression of metrics and distances can be found in [31]. The distance on the shape space is used to compare how the curves are intrinsically different. It has been seen in [38] that the simple L^2 metric on $\mathcal{M}$ given by

$$g_c^{L^2}(u,v) = \int \langle u, v \rangle \, || \, c'(t) \, || \, dt \tag{30}$$

where $\langle \cdot, \cdot \rangle$ is the Riemannian metric on $SO(n)$, induced a vanishing metric on the shape space, that is, we can not differentiate shape with this metric. To overcome this difficulty, the family of *elastic metric*, derived from the Sobolev metric [10], [23], has been investigated for it is non-vanishing on the shape space. In the case of an Euclidean space $\mathbb{R}^n$, it admits the expression:

$$g_c^{a,b}(u,v) = \int \left(a^2 \langle D_l u^N, D_l v^N \rangle + b^2 \langle D_l u^T, D_l v^T \rangle \right) || \, c'(t) \, || \, dt, \tag{31}$$

where $D_l u = h' / \, || \, c' \, ||$, $D_l u^T = \langle D_l u, w \rangle w$, with $w = c' / \, || \, c' \, ||$ and $D_l u^N = D_l u - D_l u^T$ this way, $\left(D_l u^N, D_l u^N \right)$ defines a mobile frame along the curve c, see Figure 3. Here, we are only interested in the special metric that has been proposed in [31], and which is an adaptation of the *elastic metric* for the Riemannian manifold:

$$g_c(u,v) = \langle u(0), v(0) \rangle + \int \left(\langle \nabla_l u^N, \nabla_l v^N \rangle + \frac{1}{4} \langle \nabla_l u^T, \nabla_l v^T \rangle \right) || \, (c't) \, || \, dt, \tag{32}$$

With this metric, the starting point of the curves intervenes explicitly and the metric admits a quite simple form. With the SRV framework, the length of a path of a curve (and not the length of a curve in $SO(n)$) becomes then

$$L(c) = \int_0^1 \sqrt{|| \, x(s) \, ||^2 + \int_0^1 || \, \nabla_{\partial c / \partial s} q(s,t) \, ||^2 \, dt ds}. \tag{33}$$

Once geometry has been settled in $\mathcal{M}$, the geometry of the shape space can be derived from its quotient structure. Let Before the tangent bundle be decomposed into a vertical and a horizontal subspace:

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

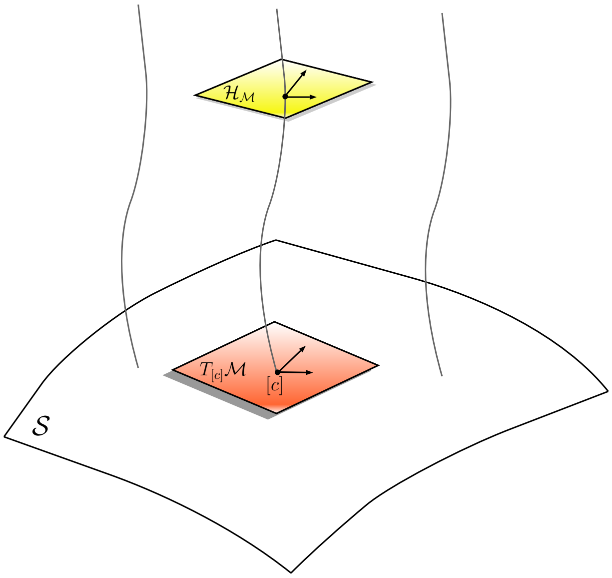


Figure 4. The tangent space $T_{[c]}\mathcal{M}$ at a point $[c]$ in the shape space $\mathcal{S}$ is isomorphic to the horizontal part $\mathcal{H}_{\mathcal{M}}$ of the tangent space at a point on the associated fiber.

$T\mathcal{M} = \mathcal{H}_{\mathcal{M}} \oplus \mathcal{V}_{\mathcal{M}}$, with $\mathcal{V}_{\mathcal{M}} = ker\,(T_c\pi)$ and T_c the tangent map, $\pi : \mathcal{M} \to \mathcal{S}$ the fiber bundle, and $\mathcal{H}_{\mathcal{M}} = (\mathcal{V}_{\mathcal{M}})^{\perp}$. This metric is reparametrisation invariant, that is, constant along the fibers, hence we have

$$g_c(u_{\mathcal{H}}, v_{\mathcal{H}}) = [g]_{\pi(c)}\,(T_c\pi(u), T_c\pi(v)) \tag{34}$$

where $[g]$ denotes the metric *on the shape space*. A similar result in a different (but still close) context is used in [56], lemma 1. In terms of distances, this can be understood in the following sense. The geodesic $s \mapsto [c](s)$ between $[c_0]$ and $[c_1]$ in the shape space is the projection of the horizontal geodesic linking c_0 to the fiber containing c_1. In fact, the horizontal geodesic between c_0 of c_1 intersects the fiber at c_1 at the reparametrised version of c_1, $c_1 \circ \phi$ which gives the distance in the shape space:

$$[d]([c_0], [c_1]) = d_g(c_0, c_1 \circ \phi) \tag{35}$$

where $[d]$ denotes the distance in $\mathcal{S}$, and d_g denotes the distance on the space of curves induced by the aforementioned Riemannian metric. In the TSRV formulation, the distance problem of eq. (35) yields an optimisation problem:

$$[d]([c_0], [c_1]) = \underset{\phi \in \mathcal{D}}{inf} \left(\int_0^1 ||\; q_0(t) - q_1(\phi(t))\sqrt{\phi'(t)}\; ||^2 \right)^{1/2}, \tag{36}$$

which is solved by a traditional gradient descent algorithm or a dynamic linear programming [13]. Finally, we have to mention that in a practical situation, the above formula has to be discretised. This is the object of [32]. Formulae are essentially similar, but in this setting, a curve is now represented by a set of points $c_{disc}(x_0, x_1, \cdots, x_n)$ and the tangent space turns into

$$T_{disc}\mathcal{M} = \{v = (v_0, v_1, \cdots, v_n), v_i \in T_{x_i}SO(n),\ \forall i\}\,. \tag{37}$$

Concerning the metric on the space of curves, it becomes

$$g_{c_{disc}}(u, v) = \langle u_0, v_0 \rangle + \frac{1}{n} \sum_{i=0}^{n-1} \langle \nabla_{\partial c/\partial s} q^u \left(0, \frac{k}{n}\right), \nabla_{\partial c/\partial s} q^v \left(0, \frac{k}{n}\right) \rangle \quad \forall u, v \in T_{disc}\mathcal{M} \tag{38}$$

where, as before, for a $u \in T_{c_{disc}}\mathcal{M}$, we define a path of piecewise geodesic curves $(s,t) \mapsto c^u(s,t)$ such that the following traditional initial conditions are fulfilled

$$\begin{aligned} c^u\left(0,\frac{k}{n}\right) &= x_k, \text{ and} \\ (\partial c^u/\partial t)\left(0,\frac{k}{n}\right) &= n\log_{x_k}(x_{k+1}). \end{aligned}$$

This is the discrete analogue of the tangent vector of a continuous curve at time t. The log function is the inverse of the exponential map on the base manifold, $SO(n)$ for us, and here $c^u\,(s,\cdot)$ must be a geodesic on $SO(n)$ between $x_{k/n}$ and $x_{(k+1)/n}$. The SRV function that appears in the formula refer to the SRV function of the piecewise geodesics $c^u\,(s,\cdot)$. Then, the discretised version of the SRV function, $q_k = \sqrt{n}\ \log_{x_k}(x_{k+1})/\sqrt{||\ \log_{x_k}(x_{k+1})\ ||}$ is such that

$$\nabla_{\partial c/\partial s}q\left(s,\frac{k}{n}\right) = \nabla_{\partial c/\partial s}q_k(s) \tag{39}$$

3.3. *The geodesic equation in the Lie group case*

Before giving the geodesic equation in the space of curves on a Lie group, we start with some preliminaries. We recall some useful facts about Lie group and Lie algebra, for those who are not familiar with these objects.
A metric $\langle\cdot,\cdot\rangle$ on a Lie group is said to be left invariant if:

$$\langle u,v\rangle_b = \langle (dL_a)_b u, (dL_a)_b v\rangle_{ab} \tag{40}$$

where $(dL_a)_b$ is the derivative in the manifold field sense (so the tangent map) of the left translation L_a at b. A left-invariant metric gives the same number whenever the vectors are translated on the left. It is straightforward to adapt this definition to a right-invariant metric. A metric that is both left and right invariant is called a bi-invariant metric. A Lie group endowed with a bi-invariant metric has plenty of import properties that can be exploited for our study of curves on shape spaces. We list some of them in the following.

- The geodesics through e (the identity element) are the integral curves $t \mapsto exp(tu),\ u \in \mathfrak{g}$, that is, the one-parameter groups. Also, because left and right are isometries and isometries maps geodesics to geodesics, the geodesics through any point $a \in G$ are the left (right) translates of the geodesics through e

$$\gamma(t) = L_a\left(exp(tu)\right),\ u \in \mathfrak{g}. \tag{41}$$

  Of course, we have

$$\gamma'(0) = (dL_a)\,e(u). \tag{42}$$

- The Levi-Civita connection is given by : $\nabla_X Y = \dfrac{1}{2}[X,Y],\ \forall X,Y \in \mathfrak{g}$
- The curvature tensor is given by : $R(u,v)w = \dfrac{1}{4}[[u,v],w]$

where $[\cdot,\cdot]$ denotes the Lie bracket. We can now link these formulas to our based manifold $SO(n)$. A Killing form, B, of a Lie algebra is a symmetric bilinear form $B : \mathfrak{g}\times\mathfrak{g} \longrightarrow \mathbb{C}$ given by $B(u,v) = tr(ad(u)\circ ad(v))$, where tr denotes the trace operator and ad denotes the adjoint representation of the group, namely, the map $ad : G \longrightarrow GL(\mathfrak{g})$ such that, for all $a \in G$ $ad_a : \mathfrak{g} \longrightarrow \mathfrak{g}$ is the *linear isomorphism* defined by $ad_a = d(R_a^{-1}\circ L_a)_e$. If we now assume B to be negative-definite, then -B is an inner product and is adjoint invariant. Thus, it is a classical result of the Lie theory that -B induces a bi-invariant metric on G. Furthermore, the Ricci curvature is given by $Ric(u,v) = -\dfrac{1}{4}B(u,v)$.

The Lie algebra of $SO(n)$ is the set of skew-symmetric matrices which verifies $M^T = -M$. The Killing form on $SO(n)$ is given by $B_{\mathfrak{so}(n)} = (n-2)tr(XY)$, and as a result of the skew symmetry, we have $-B_{\mathfrak{so}(n)} = (n-2)tr(XY^T)$. Therefore, it induces a bi-invariant metric and the previous formula can be plugged into the expression of the metric on the space of curves. To conclude these preliminaries, we see that because of the simpler form of the parallel transportation and of the metric, the distance equations (36) are now easier to handle.
It is now time to give the geodesic equation, relative to our chosen measure. As a result of the TSRV, the geodesic equation takes a much simpler form than what can be found in [31] and [32]. The formula can be found in [13]. For the sake of completeness, we give a reformulated proof in Annexe B. Recall that a geodesic is a particular path of curves. A path of curve is a continuous set of curve $s \mapsto c(s,\cdot)$ such that for each s, $c(s,\cdot)$ is a point in $\mathcal{M}$, or, equivalently, a curve in M, (see Figure ??). Thus, for each curve of the path of curves, we can defined its TSRV function. Then for all $s \in [0,1]$, we have (we omit the letter 's' for clarity): $q = \dfrac{\partial c/\partial t}{\sqrt{\| \partial c/\partial t \|}}$

Theorem 3. *A path of curves* $[0,1] \ni s \mapsto (c(s,0), q(s,t))$ *(t is the parameter of the curve $c(s,\cdot)$)is a geodesic on $\mathcal{M}$ if and only if*

$$\nabla_{\partial c/\partial s}\left(\nabla_{\partial c/\partial s} q(s,t)\right)(s,t) = 0 \quad \forall s,t \tag{43}$$

Proof. Annexe B □

Thus, we have a quite familiar expression for the geodesic interpolation between two curves c_0 and c_1, expressed in their TSRV domain:

$$F_{Lie}^{-1}\left((1-s)F_{Lie}(c_0) + sF_{Lie}(c_1)\right) \tag{44}$$

for $s \in [0,1]$. This expression is nothing but a linear interpolation on the transported tangent space.
We have almost all the ingredients now to give the procedure for nonstationary processes characterization and comparison. We first adapt the example given in [13] for curves on SO(3) to give the piecewise geodesic as follows:

$$c(t) = \sum_{k=0}^{n-1} \chi_{k,k+1}(t) exp\left((t-k) log_{c_k}(c_{k+1})\right) c_k. \tag{45}$$

Notice that this geodesic is in the base manifold, and not in the space of curves. Thus, the geodesics are expressed in terms of one-parameter groups. In order to have a curve at least $\mathcal{C}^1$, and also because the distance between the W_i matrices can be quite high, we interpolate first. There are many ways to interpolate on $SO(n)$, see [45] for example, but one of the simplest is to interpolate in the tangent space, which is Euclidean, and to go back to the manifold via the exponential map [24,45].

Finally, our procedure to compare closed curves associated with the time evolution of the spectral measure for nonstationary process is the following

1. **Input** : a set of rotation matrices $\{W_i\}_i$, seen as a partial observation of a closed trajectory on $SO(n)$.
2. Interpolate with splines between matrices W_i [24,45].
3. Go back *in the base manifold* $SO(n)$.
4. Compute the distance defined by (36).
5. **Output** : distance between two curves in the manifold defined by the set of curves in $SO(n)$.

We note that geodesic shooting [31,43] or other path straightening methods could ne applied to obtain a geodesic path *between two curves*, and between the shapes of the two curves.

 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717

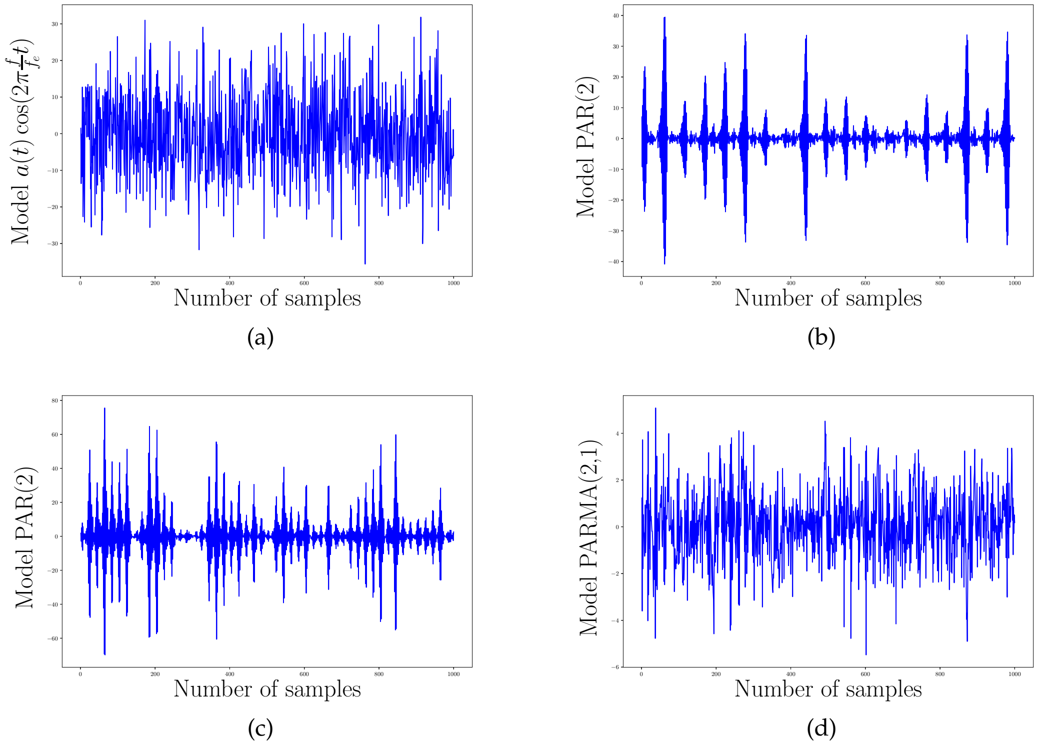


Figure 5. 1000 samples of PC processes generated by (a) a modulated zero mean and unit variance stationary random process $a(t)$, (b) a periodic AR(2) model with a period of 54 points, (c) a periodic AR(2) model with a period of 20 points, and (d) a periodic ARMA(2,1) model with a period of 20 points

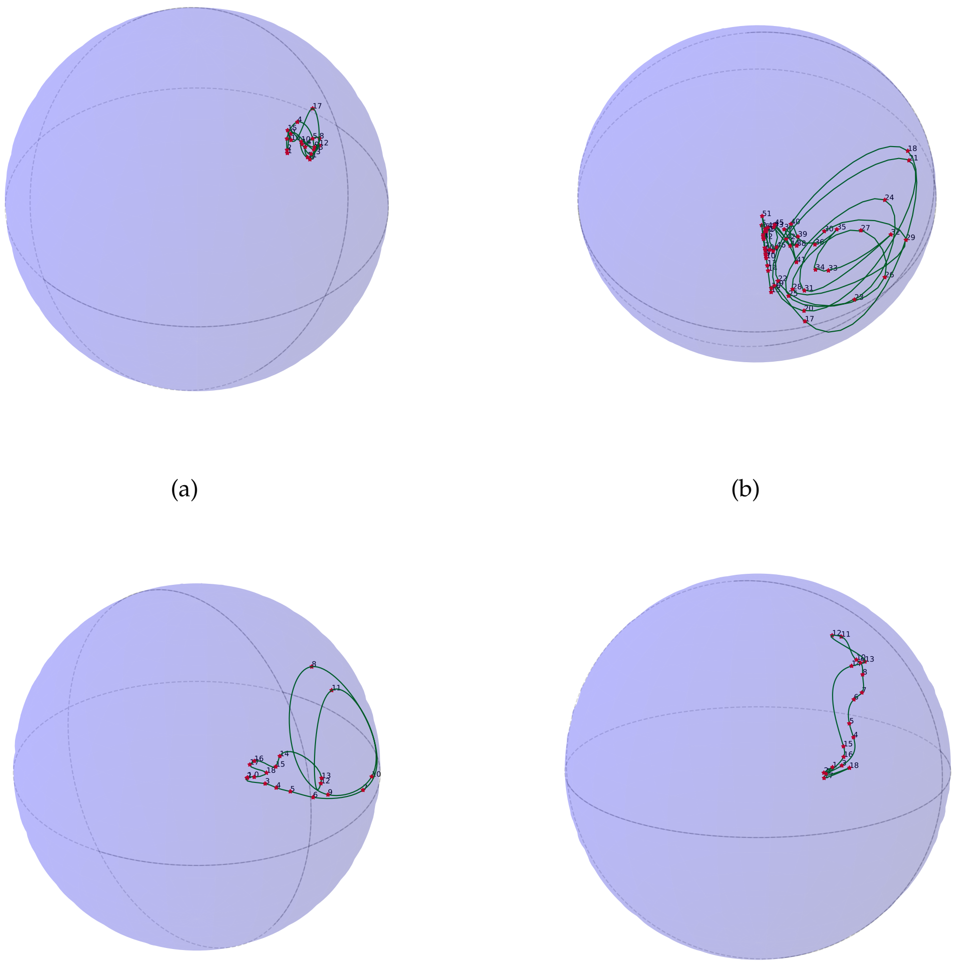
3.4. Results

In order to expose how the approach of this work gives interesting results for PC processes understanding, we propose to compare four PC processes, displayed along with Figure 5. We also bring their corresponding SO(3) representation on Figure 6. For this scenario we have generated four PC processes with 1000 samples each. A classical amplitude modulated model $a(t)\cos(2\pi f/f_e\ t)$ where $a(t)$ is a zero mean and unit variance stationary random process with a period of 20 points, a periodic AR(2) with a period of 20 points, a periodic AR(2) with a period of 54 points, and a periodic ARMA(2,1) with a period of 20 points have been generated. We have used the R package PerARMA to generate the periodic ARMA and AR signals and we finally used the PerPACF function of this package to estimate the 20 (or 54) sequences of 3 parcors each. The analysis of Figure 5 with the Figure 6 shows that the spectral measure of the amplitude modulated signal of Figure 5-(a) has dilation matrices which do not spread a lot, we could think that this process is almost stationary due to the weak distance between each matrices. A contrario, whereas the temporal form of the PARMA(2,1) signal of Figure 5-(d) is quite identical to the amplitude modulated signal of Figure 5-(a), their representation on SO(3) is very different. The spectral measure of the PARMA(2,1) signal spread much more. Lastly, when we observe the Figure 5-(b) and Figure 5-(c) which are generated with the same model but with a different period, we can see that the more the number of points per period is important, the more the curve wraps.

To end this analysis by the example, we have computed the distance defined by (36) between the PC process of Figure 7 and all the PC processes studied and displayed on Figure 5 and Figure 6. The distances are reported inside the Table. 1. Clearly, the distances between the shapes of the curves characterizing the spectral measure of each PC process, reveal some spectral proximity between the PC processes benchmarked. The PAR(2) and PARMA(2,1) are the two models which are closed to the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717



(a) (b)

(c) (d)

Figure 6. Representation inside the ball of radius π of the 4 PC processes drawn in Figure 5, arranged in the same order.

PAR(2) signal model of reference. Their spectral measure evolves consequently in a similar way with one major loop and a second less important. These observations let open besides the question of the topology of theses curves and how it could be used for the classification.

4. Conclusion

We have introduced a new vision of stochastic processes through geometry induced by dilation. The dilation matrices of given processes were obtained by a composition of rotations whose angle correspond to the well-known parcors, reflexion coefficients or Verblunski coefficients. The advantage of working with these particular matrices is that they are strongly related to the stochastic measure of the process, and thus, to its spectra. Furthermore, the dilation theory is independent of the stationarity of the underlying process; when the signal is stationary, its dilation operator is related to the Naimark dilation whereas when the signal is nonstationary, a set of dilation matrices is obtained and it is related to the Kolmogorov decomposition. Rigorously, dilation matrices are infinite dimensional, although we turn them into rotation matrices by truncation. Each of them belongs to the Special Orthogonal Group

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717

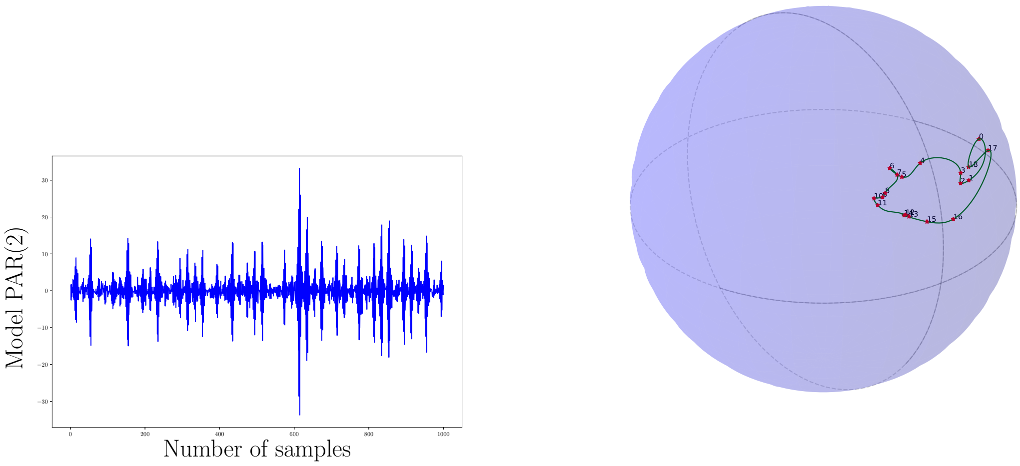


Figure 7. A PAR(2) signal with a period of 20 points, 1000 samples were generated, and its corresponding SO(3) representation inside the ball of radius π.

Table 1. Table of the distances between all the PC processes of Figure 5 to the gold standard PC process of Figure 7 through the distance of their curves' shapes on SO(3). We have applied here a DP to solve the optimization assignment problem.

Model of Signal displayed in Figure 5	Distance to the signal of Figure 7
(a)	8.97
(b)	9.11
(c)	5.12
(d)	3.92

$SO(n)$ or the Special Unitary Group $SU(n)$ depending on the real- or complex-valued process under study. We focused our attention on the Periodically Correlated (PC) class of nonstationary processes for which a timely ordered set of dilation matrices describes the process measure. This set draws a closed curve on the Lie group of rotation matrices, and describing or classifying the different PC processes is made by curves comparison. We use for that the Square Root Velocity (SRV) function which represents a curve by its starting point and by its normed velocity vector on the space or curves. The metric in the space of curve naturally extends to the space of shapes. It is then possible to compare the shape of curves when the metric is translated into the Lie algebra, achieving therefore a closed-form expression and easy computation. Nonstationary processes are then characterized via their embedded curves.

Author Contributions: Conceptualization, Mael Dugast; Formal analysis, Mael Dugast and Guillaume Bouleux; Funding acquisition, Guillaume Bouleux and Eric Marcon; Investigation, Mael Dugast; Methodology, Guillaume Bouleux; Supervision, Guillaume Bouleux and Eric Marcon; Writing – original draft, Mael Dugast; Writing – review and editing, Guillaume Bouleux.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A Defect operator, elementary rotation

Introducing the *defect operator* of a contraction T as being $D_T = (I - T^*T)^{1/2}$, we have the following factorisation:

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at Entropy 2018, 20, 717; doi:10.3390/e20090717

$$\begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} = \begin{pmatrix} X^{1/2} & 0 \\ Z^{1/2}\Gamma^* & Z^{1/2}D_\Gamma \end{pmatrix} \begin{pmatrix} X^{1/2} & \Gamma Z^{1/2} \\ 0 & D_\Gamma Z^{1/2} \end{pmatrix} \tag{A1}$$

where X and Y are positive matrices. Note that this is a Cholesky factorisation-type result. This type of decomposition is used as the square root of matrices in the construction of the dilation. A corollary is that the operator $\begin{pmatrix} I & T \\ T^* & I \end{pmatrix}$ is positive if and only if T is a contraction.

Theorem 4. *Let X and Y be operators in z. The following statements are equivalent :*

- *There exists a contraction Γ in z such that $X = \Gamma Y$,*
- *$X^*X \leqslant Y^*Y$.*

Proof. This result can be proved by taking the contraction Γ with respect to $\Gamma Xh = Yh$. [53]. □

As a corollary If, $X^*X = Y^*Y$, then there exists a partial isometry V such that $VX = Y$. It is easy to see that we can choose V to be the contraction Γ defined above. Isometry V can also be assumed unitary. For a positive operator $A \in \mathcal{L}(\mathcal{H})$, if we denote by $A^{1/2}$ its unique positive square root, then every L such that $L^*L = A$ is related to $A^{1/2}$ by $A^{1/2} = VL$ (or $A^{1/2} = L^*V^*$).
Let us state another theorem that intervene much in Constantinescu's factorisation of positive-definite kernel. Note that in the following, $\mathcal{R}(\Gamma)$ will denote the close range of the operator Γ. We first start with a basic case:

Theorem 5 (row contraction)**.** *Let $T = [T_1 \quad T_2] \in \mathcal{L}(\mathcal{H}_1 \oplus \mathcal{H}_2, \mathcal{H})$, then $|| \; T \; || \leqslant 0$ if and only if there exists contractions $\Gamma_1 \in \mathcal{L}(\mathcal{H}_1, \mathcal{H})$ and $\Gamma_2 \in \mathcal{L}(\mathcal{H}_2, \mathcal{H})$ such that*

$$T = [\Gamma_1 \quad D_{\Gamma_1^*}\Gamma_2] \tag{A2}$$

Proof. The proof is a simple application of Theorem4. For the if part, it is obvious that we can take Γ_1 to be T_1. Then $|| \; T \; || \leqslant 1$ implies

$$I - TT^* = I - \Gamma_1\Gamma_1^* - T_2T_2^* \geqslant 0 \tag{A3}$$

with $D^2_{\Gamma_1^*} \geqslant T_2T_2^*$. Hence, there exists Δ such that $\Delta D_{\Gamma_1^*} = T_2^*$. Choosing $\Gamma_2 = \Delta^*$ finishes the argument. □

In the same way as that of the Cholesky factorisation, we can write down the defect operator for the whole contraction $T = [T_1 \quad T_2]$ [53] to be

$$D_T^2 = \begin{pmatrix} D_{\Gamma_1} & 0 \\ -\Gamma_2^*\Gamma_1 & D_{\Gamma_1} \end{pmatrix} \begin{pmatrix} D_{\Gamma_1} & -\Gamma_1^*\Gamma_2 \\ 0 & D_{\Gamma_1}. \end{pmatrix} \tag{A4}$$

Therefore, with Theorem 5, we have an operator α such that

$$D_T = \begin{pmatrix} D_{\Gamma_1} & 0 \\ -\Gamma_2^*\Gamma_1 & D_{\Gamma_1} \end{pmatrix} \alpha \tag{A5}$$

Similarly,

$$D^2_{T^*} = (D_{\Gamma_1^*} D_{\Gamma_2^*} D_{\Gamma_2^*} D_{\Gamma_1^*}) \tag{A6}$$

and the general case is

Theorem 6 (Structure of row contraction)**.** *The following are equivalent :*

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717

- *The operator* $T^n = [T_1 \quad T_2 \quad \cdots T^n]$ *in* $\mathcal{L}(\oplus_{k=1}^n \mathcal{H}_k, \mathcal{H}')$ *is a contraction*
- $T_1 = \Gamma_1$ *is a contraction and, for* $k > 2$, *there exists uniquely determined contractions* $\Gamma_k \in \mathcal{L}(\mathcal{H}_k, \mathcal{R}(\gamma_k))$ *such that* $T_k = D_{\Gamma_1^*} D_{\Gamma_2^*} \cdots D_{\Gamma_{k-1}^*} \Gamma_k$.

Furthermore, the defect operators of the whole contraction T are of the form

$$D_T^2 = \begin{pmatrix} D_{\Gamma_1} & 0 & \cdots & 0 \\ -\Gamma_2^*\Gamma_1 & D_{\Gamma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\Gamma_n^* D_{\Gamma_{n-1}^*} \cdots D_{\Gamma_2^*} & -\Gamma_n^* D_{\Gamma_{n-1}^*} \cdots D_{\Gamma_3^*}\Gamma_2 & \cdots & D_{\Gamma_n} \end{pmatrix} \begin{pmatrix} D_{\Gamma_1} & -\Gamma_1^*\Gamma_2 & \cdots & -\Gamma_1^* D_{\Gamma_2^*} \cdots D_{\Gamma_{n-1}^*}\Gamma_n \\ 0 & D_{\Gamma_2} & \cdots & -\Gamma_2^* D_{\Gamma_3^*} \cdots D_{\Gamma_{n-1}^*}\Gamma_n \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{\Gamma_n} \end{pmatrix} \tag{A7}$$

and

$$D_{T^*}^2 = D_{\Gamma_1^*} \cdots D_{\Gamma_n^*} D_{\Gamma_n^*} \cdots D_{\Gamma_1^*} \tag{A8}$$

Proof. It can be proved straightforwardly by induction. □

This construction permits to understand the apparition of the operators α and β in the publications of Constantinescu which are used to identify the defect space of the components (the underlying contractions of a row contraction) of a row contraction with the defect space of the row contraction itself. Same results are readily obtained for a column contraction of the form $T = \begin{pmatrix} T_1 \\ \vdots \\ T_2 \end{pmatrix}$.

Appendix B Geodesic equation in the space of curve $\mathcal{M}$

To have a complete insight on the geodesic equation, we give the proof for a more general case that arises when considering the SRV and not only the TSRV function of a curve, that is, the curves are parametrised by their starting point and their velocity, but their starting points are not transported to the identity.

Theorem 7. *A path of curves* $[0,1] \ni s \mapsto (c(s,0), q(s,t))$ *(t is the parameter of the curve $c(s,\cdot)$)is a geodesic on $\mathcal{M}$ if and only if:*

$$\nabla_{\partial c/\partial s} c(s,0) + \int_0^1 \mathcal{R}\left(q(s,t), \nabla_{\partial c/\partial s} q(s,t)\right)(c(s,0))dt = 0 \quad \forall s \tag{A9}$$

$$\nabla_{\partial c/\partial s}\left(\nabla_{\partial c/\partial s} q(s,t)\right)(s,t) = 0 \quad \forall s,t \tag{A10}$$

Similarly to [31] and [57], we consider a variation of the path $s \mapsto c(s,0), q(s,t)$ starting and ending at the same points, we denote $\{(c(s,0,\tau), q(s,t,\tau))\}$. In Figure (A1), to get a clear picture, we have represented a variation of a path of curves with fixed starting and ending points. Although similar, the situation here is a bit different because of the representation of the curve through its SRV function, which we can hardly represent. However, the process remains similar. We emphasise the subtle difference with [31]. Here, we work directly in the tangent space representation, via the SRV representation, and not with "the whole family" of curves $c(s,t,\tau)$. We denote $\partial_\tau c(s,0,\tau) = \dfrac{\partial c(s,0,\tau)}{\partial \tau}$, and similarly for $\partial_s c(s,0,\tau)$ and $\partial_\tau c(s,0,\tau)$. The energy of the path indexed by τ is

$$E(\tau) = \frac{1}{2}\int_0^1 \langle \partial_s c(s,0,\tau), \partial_s c(s,0,\tau)\rangle + \langle \nabla_{\partial c/\partial s} q(s,t,\tau), \nabla_{\partial c/\partial s} q(s,t,\tau)\rangle ds. \tag{A11}$$

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 23 July 2018 doi:10.20944/preprints201807.0406.v1

Peer-reviewed version available at *Entropy* **2018**, *20*, 717; doi:10.3390/e20090717

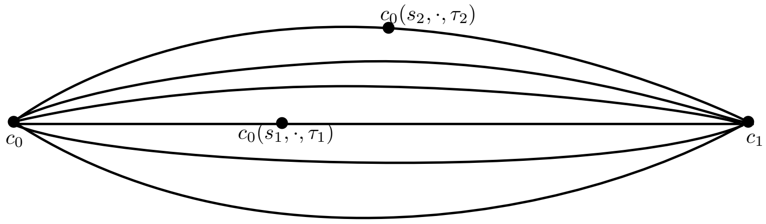


Figure A1. we consider a beam of curves, which consists in a slight modification of the geodesic. The different curves are indexed by τ. The idea is to find which of these curves gives the minimal energy to go from c_0 to c_1.

Recall that the derivative of the inner product is given by $\frac{d}{dx}\langle f(x), f(x)\rangle = 2 * \langle f(x), \frac{df}{dx}\rangle$. Then

$$E'(0) = \int_0^1 \langle \nabla_{\partial c/\partial\tau} \frac{\partial c}{\partial s}(s,0,0), \frac{\partial c}{\partial s}(s,0,0)\rangle + \langle \nabla_{\partial c/\partial\tau}\nabla_{\partial c/\partial s} q(s,t,0), \nabla_{\partial c/\partial s} q(s,t,0)\rangle ds \tag{A12}$$

with $\nabla_{\partial c/\partial s}(\partial_\tau c(s,0,\tau)) = \nabla_{\partial c/\partial\tau}(\partial_s c(s,0,\tau))$ and owing to the curvature tensor $\mathcal{R}(\partial_\tau c(s,0,\tau), \partial_s c(s,0,\tau))(q(s,t,\tau) = \nabla_{\partial c/\partial\tau}\nabla_{\partial c/\partial s}(q(s,t,\tau)) - \nabla_{\partial c/\partial s}\nabla_{\partial c/\partial\tau}(q(s,t,\tau))$ we have

$$\begin{aligned} E'(0) =& \\ &\int_0^1 \langle \nabla_{\partial_s c}\partial_\tau c(s,0,\tau), \partial_s c(s,0,\tau)\rangle + \langle \mathcal{R}(\partial_\tau c(s,0,\tau), \partial_s c(s,0,\tau))\, q(s,t,\tau), \nabla_{\partial_s} q(s,t,\tau)\rangle \\ &+ \langle \nabla_{\partial_s c}\nabla_{\partial_\tau c} q(s,t,0), \nabla_{\partial_s c} q(s,t,0)\rangle\, ds. \end{aligned} \tag{A13}$$

Integrating by parts now, allows to have

$$\int_0^1 \langle \nabla_{\partial_\tau c}\partial_s c(s,0,\tau), \partial_s c(s,0,\tau)\rangle ds = -\int_0^1 \langle \nabla_{\partial_s c}\partial_s c(s,0,\tau), \partial_\tau c(s,0,\tau)\rangle ds$$
$$\int_0^1 \langle \nabla_{\partial_s c}\nabla_{\partial_\tau c}(q(s,t,\tau)), \nabla_{\partial_s} q(s,t,\tau)\rangle = -\int_0^1 \langle \nabla_{\partial_s c}\nabla_{\partial_s c}(q(s,t,\tau)), \nabla_{\partial_\tau} q(s,t,\tau)\rangle$$

which yields to

$$\begin{aligned} E'(0) = \int_0^1 &(-\langle \nabla_{\partial_s c}\partial_s c(s,0,\tau), \partial_\tau c(s,0,\tau)\rangle) \\ &+ \langle \mathcal{R}(\partial_\tau c(s,0,\tau), \partial_s c(s,0,\tau))\, q(s,t,\tau), \nabla_{\partial_s} q(s,t,\tau)\rangle \\ &+ (-\langle \nabla_{\partial_s c}\nabla_{\partial_s c} q(s,t,0), \nabla_{\partial_\tau c} q(s,t,0)\rangle) ds, \end{aligned} \tag{A14}$$

for any vector fields X, Y, Z, W, $\langle \mathcal{R}(X,Y)Z, W\rangle = -\langle \mathcal{R}(W), Z\rangle$, we consequently obtain

$$\begin{aligned} E'(0) = -\int_0^1 &\langle \nabla_{\partial_s c}\partial_\tau c(s,0,\tau), \partial_s c(s,0,\tau)\rangle \\ &+ \langle \mathcal{R}(q(s,t,\tau), \nabla_{\partial_s} q(s,t,\tau))(\partial_s c(s,0,\tau)), \partial_\tau c(s,0,\tau\rangle \\ &+ \langle \nabla_{\partial c/\partial s}\nabla_{\partial c/\partial\tau} q(s,t,0), \nabla_{\partial c/\partial s} q(s,t,0)\rangle ds. \end{aligned} \tag{A15}$$

Geodesic corresponds to minimal energy. It means that every other path that starts and ends at the same points should require more energy to travel than the geodesic. We then have to solve $E'(0) = 0$ for every $\partial_\tau c(s,0,\tau)$ and every $\nabla_{\partial_\tau}(q(s,t,\tau))$. This gives the result.
Now when the framework is given by the TSRV and not by the SRV, only the second part of the geodesic equation remains as a result of the fixed starting point which corresponds to the identity element. This very much simplifies the equation, even though the derivation is the same.

References

1. G. Ammar, W. Gragg, and L. Reichel, *Constructing a Unitary Hessenberg Matrix from Spectral Data*, no. 70 in NATO ASI Series, Springer Berlin Heidelberg, 1991, pp. 385–395.
2. M. Arnaudon, F. Barbaresco, and Le Yang, *Riemannian Medians and Means With Applications to Radar Signal Processing*, IEEE Journal of Selected Topics in Signal Processing, 7 (2013), pp. 595–604.
3. G. Arsene and T. Constantinescu, *The structure of the Naimark dilation and Gaussian stationary processes*, oper. theory, 8 (1985), pp. 181–204, https://doi.org/10.1007/BF01202813.
4. F. Auger, P. Flandrin, Y.-T. Lin, S. McLaughlin, S. Meignen, T. Oberlin, and H.-T. Wu, *Time-Frequency Reassignment and Synchrosqueezing: An Overview*, IEEE Signal Process. Mag., 30 (2013), pp. 32–41.
5. B. Balaji, F. Barbaresco, and A. Decurninge, *Information geometry and estimation of Toeplitz covariance matrices*, in 2014 Int. Radar Conf., IEEE, oct 2014, pp. 1–4.
6. F. Barbaresco, *Interactions between symmetric cone and information geometries: Bruhat-tits and siegel spaces models for high resolution autoregressive doppler imagery*, in Emerging Trends in Visual Computing, Springer, 2009, pp. 124–163.
7. F. Barbaresco, *Geometric Radar Processing based on Fréchet distance: Information geometry versus Optimal Transport Theory*, in 2011 12th Int. Radar Symp., 2011, pp. 663–668.
8. F. Barbaresco, *Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median*, in Matrix Inf. Geom., Springer Berlin Heidelberg, Berlin, Heidelberg, 2013, pp. 199–255.
9. M. Bauer, M. Bruveris, S. Marsland, and P. W. Michor, *Constructing reparameterization invariant metrics on spaces of plane curves*, Differential Geometry and its Applications, 34 (2014), pp. 139–165.
10. M. Bauer, M. Bruveris, and P. W. Michor, *Why Use Sobolev Metrics on the Space of Curves*, in Riemannian Computing in Computer Vision, Springer, Cham, 2016, pp. 233–255.
11. N. H. Bingham, *Szego's Theorem and its Probabilistic Descendants*, (2011), http://arxiv.org/abs/1108.0368.
12. S. Bochner, *Vorlesungen über fouriersche integrale*, Mathematik und ihre Anwendungen, Akademie-Verlag, (1932).
13. E. Celledoni, M. Eslitzbichler, and A. Schmeding, *Shape analysis on Lie groups with applications in computer animation*, arXiv preprint arXiv:1506.00783, (2015).
14. E. Chevallier, A. Chevallier, and J. Angulo, *Computing histogram of tensor images using orthogonal series density estimation and riemannian metrics*, in 2014 22nd International Conference on Pattern Recognition, 2014, pp. 900–905.
15. T. Constantinescu, *Schur Parameters, Factorization and Dilation Problems*, Birkhäuser Basel, Basel, 1995.
16. F. Desbouvries, *Geometrical aspects of linear prediction algorithms*.
17. F. Desbouvries, *Unitary Hessenberg and state-space model based methods for the harmonic retrieval problem*, IEE Proc. - Radar, Sonar Navig., 143 (1996), p. 346.
18. F. Desbouvries, *Non-Euclidean geometrical aspects of the Schur and Levinson-Szego algorithms*, IEEE Trans. on Information theory, 49 (2003), p. 346.
19. F. Desbouvries, *Non-euclidean geometrical aspects of the schur and levinson-szego algorithms*, IEEE Transactions on Information Theory, 49 (2003), pp. 1992–2003.
20. S. Dégerine and S. Lambert-Lacroix, *Characterization of the partial autocorrelation function of nonstationary time series*, Journal of Multivariate Analysis, 87 (2003), pp. 46–59.
21. P. Flandrin, in Time-Frequency/Time-Scale Analysis, vol. 10 of Wavelet Analysis and Its Applications, Academic Press, 1999.
22. C. Foias and A. E. Frazho, *A Geometric Approach to Positive Definite Sequences*, in The Commutant Lifting Approach to Interpolation Problems, no. 44 in OT 44 Operator Theory: Advances and Applications, Birkhäuser Basel, 1990, pp. 497–546.
23. A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin, *Bayesian data analysis*, vol. 2, CRC press Boca Raton, FL, 2014.
24. M. Hofer and H. Pottmann, *Energy-minimizing Splines in Manifolds*, ACM Trans. Graph., August 2004, vol. 23, no. 3, 2004, pp. 284–293.

25. H. L. HURD AND A. MIAMEE, *Periodically Correlated Random Sequences: Spectral Theory and Practice*, Wiley Series in Probability and Statistics, Wiley, 2007.
26. X. JIANG, L. NING, AND T. GEORGIOU, *Distances and Riemannian Metrics for Multivariate Spectral Densities*, IEEE Trans. Automat. Contr., 57 (2012), pp. 1723–1735.
27. T. KAILATH AND A. M. BRUCKSTEIN, *Naimark Dilations, State-Space Generators and Transmission Lines*, in Advances in Invariant Subspaces and Other Results of Operator Theory, R. G. Douglas, C. M. Pearcy, B. Sz.-Nagy, F.-H. Vasilescu, D. Voiculescu, and G. Arsene, eds., no. 17 in Operator Theory: Advances and Applications, Birkhäuser Basel, 1986, pp. 173–186.
28. E. KALUNGA, S. CHEVALLIER, Q. BARTHÉLEMY, K. DJOUANI, Y. HAMAM, AND E. MONACELLI, *From Euclidean to Riemannian Means: Information Geometry for SSVEP Classification*, Springer, Cham, oct 2015, pp. 595–604.
29. S. KOLOURI, S. PARK, M. THORPE, D. SLEPCEV, AND G. ROHDE, *Optimal Mass Transport: Signal processing and machine-learning applications*, IEEE Signal Process. Mag., 34 (2017), pp. 43–59.
30. S. LAMBERT-LACROIX, *Extension of Autocovariance Coefficients Sequence for Periodically Correlated Processes*, Journal of Time Series Analysis, 26 (2005), pp. 423–435.
31. A. LE BRIGANT, *Computing distances and geodesics between manifold-valued curves in the SRV framework*, arXiv preprint arXiv:1601.02358, (2016).
32. A. LE BRIGANT, *A discrete framework to find the optimal matching between manifold-valued curves*, arXiv preprint arXiv:1703.05107, (2017).
33. A. MAKAGON AND H. SALEHI, *Notes on infinite dimensional stationary sequences*, in Probability Theory on Vector Spaces IV, S. Cambanis and A. Weron, eds., vol. 1391 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1989, pp. 200–238.
34. J. MANTON, *A Primer on Stochastic Differential Geometry for Signal Processing*, IEEE J. Sel. Top. Signal Process., 7 (2013), pp. 681–699.
35. P. MASANI, *Dilations as Propagators of Hilbertian Varieties*, SIAM J. Math. Anal., 9 (1978), pp. 414–456.
36. A. MIAMEE, *Spectral dilation of L(B,H)-valued measures and its application to stationary dilation for Banach space valued processes*, Indiana Univ. Math. Journal, 38 (1989), pp. 841–860.
37. A. MIAMEE, *Periodically Correlated Processes and Their Stationary Dilations*, SIAM Journal on Applied Mathematics, 50 (1990), pp. 1194–1199.
38. P. W. MICHOR AND D. MUMFORD, *Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms*, Doc. Math, 10 (2005), pp. 217–245.
39. N. LIPENG, J. XIANHUA, AND T. GEORGIOU, *Geometric methods for structured covariance estimation*, in 2012 Am. Control Conf., IEEE, jun 2012, pp. 1877–1882.
40. A. NAPOLITANO, *Cyclostationarity: New trends and applications*, Signal Processing, 120 (2016), pp. 385–408.
41. X. PENNEC, *Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements*, Journal of Mathematical Imaging and Vision, 25 (2006), pp. 127–154.
42. X. PENNEC, P. FILLARD, AND N. AYACHE, *A Riemannian framework for tensor computing*, International Journal of Computer Vision, 66 (2006), pp. 41–66.
43. M. PILTÉ AND F. BARBARESCO, *Tracking quality monitoring based on information geometry and geodesic shooting*, 17th Int. Radar Symp., 2016, pp. 1–6.
44. A. H. SAYED, T. CONSTANTINESCU, AND T. KAILATH, *Recursive Construction of Multichannel Transmission Lines with a Maximum Entropy Property*, in Codes, Graphs, and Systems, Springer, 2002, pp. 259–290.
45. T. SHINGEL, *Interpolation in special orthogonal groups*, IMA Journal of Num. Ana., vol. 29,no. 3, July 2009, pp. 731–745.
46. B. SIMON, *CMV matrices: Five years after*, Journal of Computational and Applied Mathematics, 208 (2007), pp. 120–154.
47. B. SIMON, *Orthogonal Polynomials on the Unit Circle Part1 and Part 2*, vol. 54, American Mathematical Society, 2009.
48. S. T. SMITH, *Covariance, subspace, and intrinsic Cramér-Rao bounds*, IEEE Trans. on Signal Process., 53 (2005), pp. 1610–1630.
49. A. SRIVASTAVA, E. KLASSEN, S. H. JOSHI, AND I. H. JERMYN, *Shape Analysis of Elastic Curves in Euclidean Spaces*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 33 (2011), pp. 1415–1428.
50. K. SUN AND S. MARCHAND-MAILLET, *An information geometry of statistical manifold learning*, 2014.

51. B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy, *Harmonic Analysis of Operators on Hilbert Space*, Springer New York, New York, NY, 2010.
52. D. Timotin, *Prediction theory and choice sequences: an alternate approach*, in Advances in invariant subspaces and other results of operator theory, Springer, 1986, pp. 341–352.
53. M. C. Tseng, *Contractions, Matrix Paramatrizations, and Quantum Information*, arXiv:quant-ph/0610259, (2006).
54. M. C. Tseng and V. Ramakrishna, *Dilation Theoretic Parametrizations of Positive Matrices with Applications to Quantum Information*, arXiv:quant-ph/0610021, (2006).
55. L. Yang, M. Arnaudon, and F. Barbaresco, *Riemannian median, geometry of covariance matrices and radar target detection*, in Radar Conference (EuRAD), 2010 European, Sept. 2010, pp. 415–418.
56. Yili L. and K. M. Wong, *Riemannian Distances for Signal Classification by Power Spectral Density*, IEEE Journal of Selected Topics in Signal Processing, 7 (2013), pp. 655–669.
57. Z. Zhang, J. Su, E. Klassen, H. Le, and A. Srivastava, *Video-based action recognition using rate-invariant analysis of covariance trajectories*, arXiv preprint arXiv:1503.06699, (2015).