

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@insa-lyon.fr (M.D.); eric.marcon@insa-lyon.fr (E.M.)

* Correspondence: guillaume.bouleux@insa-lyon.fr; Tel.: +33-477-448-920

Received: 21 July 2018; Accepted: 7 September 2018; Published: 19 September 2018



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 random processes' measures.

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 four or five decades now by time-scale/time-frequency analysis [1,2], by Fourier-like representation when the processes belong to the periodically correlated (PC) subclass [3,4], or by partial correlation coefficients (parcors) series [5,6], 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 [7,8]. They consequently appear explicitly in the Orthogonal Polynomial on the Real Line/Unit Circle decomposition of the measure [9,10], on the Matrices Orthogonal Polynomials on the Unit Circle [11] and its applications [12], are the elements for the construction of dilation matrices that appear in the Cantero Moral, and Velazquez (CMV)/Geronimus, Gragg, and Teplyaev (GGT) matrices [13], for the Schur flows problem with upper Hessenberg matrices [14] that are also seen in the literature as evolution operators [10] or shift operator [15], and finally appear in the state-space representation [7,16]. The dilation theory takes its roots from the operator theory [17], which bridges the process's measure and unitary operators. In its simplest version, the dilation theory corresponds to Naimark dilation [17,18], and states that given a sequence of correlation coefficients, there exists a unitary matrix W such that

$R_n \triangleq (1\ 0\ 0\ \dots)W^n(1\ 0\ 0\ \dots)^T$ 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 [16] and the previous expression becomes $R_{i,j} \triangleq (1\ 0\ 0\ \dots)W_{i+1}W_{i+2}\dots W_j(1\ 0\ 0\ \dots)^T$. The matrices W_i are theoretically understood as infinite rotation matrices, which become finite when the correlation coefficients sequence is itself 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's measure is totally described by the set of W_i . As a consequence, the measure of the process is beautifully characterized 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 is periodic, as well as the sequence of dilation matrices, which yields a closed path as illustrated in Figure 1. This raises the question of comparison of processes by means of their dilation matrices. Many efforts have been made in the last decade to exploit the hyperbolic geometric structure not of the correlation matrices directly but of the related parcors when obtained in stationary conditions [8,19–22], as well as the information geometry structure that is closely related to it [23–25]. 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. Treatment of the nonstationary case has not been tackled to our knowledge with the previously mentioned approaches. In this paper, we hope to initiate interest in filling this gap by extending the representation and the characterization of the processes's measure in a nonstationary context, inspired on the one hand by information geometric application and interpretation of parcors [19,20] or correlation matrices [26–28], and on the other hand by theories and applications dealing with curves on manifolds [29,30], closely related to some aspects of Euler's equations [31]. Therefore, we combine Constantinescu's approach to dilation and shape analysis to the propose of seeing stochastic processes as elements of a Lie group. Characterizing the time-varying measure of the process is now tackled by studying curves (or sampled curves) on special groups.

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 [16,18] or the lattice filter structure approach [32,33] are also discussed and the strong connection between parcors and the dilation matrices is emphasized. 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 Appendices A and B.

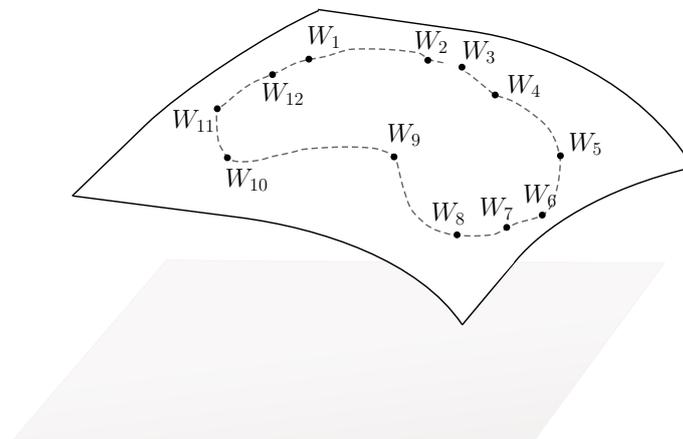


Figure 1. Illustration of a sampled closed trajectory drawn in $SO(n)$ or $SU(n)$ that materializes the time varying of the Periodically Correlated (PC) measure for a stochastic process. Each W_i is a dilation matrix built through the parcors. Recall that a PC process is a process such that $R_{s,t} = R_{s+T,t+T}$ for a certain T , where $R_{\cdot,\cdot}$ stands for the correlation function of the process.

2. The Structure of Semi-Positive-Definite Matrices and the Dilation Theory

2.1. Outline of the Dilation Theory

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 another application \mathcal{F} such that f is the orthogonal projection of \mathcal{F} :

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

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, we obtain a unitary dilation. This approach has been for example developed in [34–36] 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 following unitary operator

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

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

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

In other words, the elementary rotation of a contraction also corresponds to the unitary dilation operator of the contraction. This operator is called the Julia operator, and corresponds to the Halmos extension [15] 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. A sequence of operators $\{R_n\}_{n=1}^{\infty}$ acting on \mathcal{H} is said to be positive if

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

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

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

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 [37]. Owing to this property, we obtain the existence of an operator U such that [18,38] and Theorem 1.1 in [39]:

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

as a result of the Naimark dilation theorem [17]. Furthermore, if $\mathcal{K} = \bigvee_{n \geq 0} U^n \mathcal{H}$, where \bigvee denotes the linear span, then U is unique up to an isomorphism.

2.1.3. Dilation and Measure

From Bochner’s theorem [37], we know 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 [16]. 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

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

is positive-definite. This 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 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 $R_{i,j} = \left(1 \ 0 \ 0 \ \dots \ 0\right) W_i W_{i+1} \dots W_{j-1} \left(1 \ 0 \ 0 \ \dots \ 0\right)^T$. Let the general framework where $R_{i,j}$ is a complex operator satisfying $R_{i,j} \in \mathcal{L}(\mathcal{H}_j, \mathcal{H}_i)$ with $\{\mathcal{H}_n\}_n$ a sequence of Hilbert spaces and \mathcal{L} the set of linear applications. For example, consider the stochastic process $\{X_n\}_n$, where $X_n \in L^2(\mathbb{P})$ is a squared integrable random variable with respect to the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. Then the stochastic process can be viewed as an operator: $\tilde{X}_n : \mathbb{C} \rightarrow L^2(\mathbb{P}), \tilde{X}_n \lambda = \lambda X_n$, and the correlation kernel becomes $R_{i,j} = \tilde{X}_i^* \tilde{X}_j$.

To start the construction, let first the following theorem

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
- 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. Appendix A □

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

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

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 Appendix A. For example, we have $R_{1,2} = R_{1,1}^{1/2} \Gamma_{1,2} R_{1,1}^{1/2}$. 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 [16], we will call these contractions the Schur-Constantinescu parameters. We consider now unit variance and arbitrary size $n \times n$ for the SPD matrix, which allows us to write the correspondence as:

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

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 $G_{\Gamma_{i,i+1}}$ denotes the Givens rotation of $\Gamma_{i,i+1}$ as follows:

$$G(\Gamma_{i,i+1}) = I \oplus \begin{pmatrix} \Gamma_{i,i+1} & D_{\Gamma_{i,i+1}^*} \\ D_{\Gamma_{i,i+1}} & -\Gamma_{i,i+1}^* \end{pmatrix} \oplus I = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & & \cdots & 0 \\ & & \ddots & & & & & 0 \\ 0 & \cdots & 0 & \Gamma_{i,i+1} & D_{\Gamma_{i,i+1}^*} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & D_{\Gamma_{i,i+1}} & -\Gamma_{i,i+1}^* & 0 & \cdots & 0 \\ 0 & \cdots & & 0 & 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & & & \ddots & \vdots \\ 0 & \cdots & & & & & & & 1 \end{pmatrix} \tag{14}$$

where the “non-identity” part, consisting of a Julia operator $\left(D_{\Gamma_{i,i+1}} = \left(I - \Gamma_{i,i+1}^* \Gamma_{i,i+1} \right)^{1/2} \right)$ is located at the entry (i, i) . When the SPD matrix is Toeplitz, which corresponds to a stationary underlying process, then all dilation matrices W_i are identical and take the form

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

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

- $R_{kk} \geq 0$ for all k
- there exists a family $\{\Gamma_{k,j} \mid k, j = 1, \dots, n, k \leq j\}$ of contraction such that

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

where $B_{k,k}$ is the Cholesky's square-root of $R_{k,k}$.

and

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

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 $\{\Gamma_{m,j} \mid m = j - 1, \dots, k\}$, and finally

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

Proof. This theorem is proved in ([16], Theorem 5.3). \square

A different approach leading to the same results can be found in [40], using directly the Kolmogorov decomposition. In [32] the Naimark dilation is constructed using the lattice filter and finally applications of this decomposition in quantum mechanics are to be found in [41,42] for example.

We now give some remarks to conclude this part:

- If $R^{(n)}$ is a semi-positive definite complex-valued Toeplitz kernel, then all the $\{\Gamma_n\}$ are complex-valued and respect $|\Gamma_i| < 1$. The structure and the construction procedure for obtaining such a complex-valued parameter is identical whether the kernel is real or complex.
- The framework proposed by Constantinescu and recalled previously is quite general and can be extended to Matrix Orthogonal Polynomial on the Unit Circle (MOPUC) development. By referring to ([43], Section 3.1), matrix polynomials stem from a Szëgo recursion akin to the scalar case and thus provide a sequence of Verblunsky coefficients, or parcors that are matrices. Again in ([43], Section 3.11), a correspondence between MOPUC and CMV matrices [13] is provided, which are equivalent to dilation matrices. The construction procedure remains the same for the dilation matrices, but the parcors become in that case matrices (they are matrix-valued Verblunsky coefficients), which can be obtained by a matrix version of the Schur/Geronimus algorithm [44].

Dilation matrices being now fully introduced, we focus the attention of the reader on the hidden information contained in their timely geometrical dissemination.

3. Analysis of Curves on a Manifold Induced by the Dilation

Parcours, composing dilation matrices, have already been given a geometrical point of view, as, for example, in [8] 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 characterize individual parcors is then proposed and discussed. In [45], 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 materializes 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 [45,46], Celledoni et al. [47] and Zhang et al. [48], we propose then to give 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 emphasize 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 realizations 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, but many results have been proposed recently to overcome this difficulty and to propose closed-form formulations [30,49]. In particular, it is predicated to extract the shape of the trajectory for it contains the essential information, in a topologic sense.

To allow the curves comparison, we have based our development on the works of Le Brigant [45] and Celledoni et al. [47]. 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 principal 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. Preliminaries on Lie Groups

As we are going to deal with curves on a Lie group, we start with some preliminaries.

A metric $\langle \cdot, \cdot \rangle$ on a Lie group is said to be left invariant if:

$$\langle u, v \rangle_b = \langle (dL_a)_b u, (dL_a)_b v \rangle_{ab} \quad (20)$$

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 [50].

- 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. In addition, 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

$$\gamma(t) = L_a(\exp(tu)), \quad u \in \mathfrak{g}. \quad (21)$$

Of course, we have

$$\gamma'(0) = (dL_a) e(u). \tag{22}$$

- The Levi-Civita connection is given by: $\nabla_X Y = \frac{1}{2}[X, Y], \forall X, Y \in \mathfrak{g}$

where $[\cdot, \cdot]$ denotes the Lie bracket. We can now link these formulas to our base manifold $SO(n)$. The Killing form, B , of a Lie algebra is the symmetric bilinear form $B : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$ given by $B(u, v) = \text{tr}(ad(u) \circ ad(v))$, where tr denotes the trace operator and ad denotes the adjoint representation of the group, namely, the map $ad : G \rightarrow GL(\mathfrak{g})$ such that, for all $a \in G$ $ad_a : \mathfrak{g} \rightarrow \mathfrak{g}$ 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 compact semi-simple Lie theory that $-B$ induces a bi-invariant metric on G .

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 $B_{so(n)} = (n - 2)\text{tr}(XY)$, and as a result of the skew symmetry, we have $-B_{so(n)} = (n - 2)\text{tr}(XY^T)$. 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. In the sequel, the manifold that supports the curves is $SO(n)$ endowed with its bi-invariant metric.

3.2. Basic Outline of Geometry

Curves of interest are those living in the Lie group of real rotation matrices; this yields $c : [0, 1] \rightarrow SO(n)$. For the sake of clarity, assume that $c \in C^\infty([0, 1], SO(n))$, 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., $\mathcal{M} = \{c \in C^\infty([0, 1], SO(n)) : c'(t) \neq 0 \forall t\}$, this is in fact a sub-manifold of $C^\infty([0, 1], SO(n))$. A curve c is thus a particular point in \mathcal{M} . The tangent space at a curve c is given by

$$T_c \mathcal{M} = \left\{ v \in C^\infty([0, 1], TSO(n)) : v(t) \in T_{c(t)} SO(n) \right\} \tag{23}$$

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 $SO(n)$ (see Theorem 5.6 in [51]). When comparing two curves, it is natural that the distance between these two curves should remain the same if the curves are only reparametrized, 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 discretized as we will see in the sequel, doing a reparametrization is equivalent to changing the chosen points (see Figure 2). A reparametrization is represented by increasing diffeomorphism $\phi \in \mathcal{D} : [0, 1] \rightarrow [0, 1]$ 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:

$$g_{c \circ \phi}(u \circ \phi, v \circ \phi) = g_c(u, v) \tag{24}$$

for all $c \in \mathcal{M}, u, v \in T_c \mathcal{M}$ and $\phi \in \mathcal{D}$.

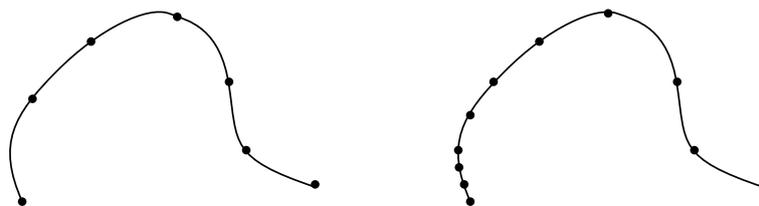


Figure 2. Example of a reparametrization of a curve. Here, it consists in changing the discretization with nonlinear time sample.

This property is called reparametrization invariance. We insist 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

$$d_{\mathcal{M}}(c_0 \circ \phi, c_1 \circ \phi) = d_{\mathcal{M}}(c_0, c_1) \tag{25}$$

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

$$c_0 \sim c_1 \iff \exists \phi \in \mathcal{D} : c_0 = c_1 \circ \phi. \tag{26}$$

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:

$$\mathcal{S} = \mathcal{M} / \sim, \text{ or } \mathcal{S} = \mathcal{M} / \mathcal{D}. \tag{27}$$

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

$$d_{\mathcal{S}}([c_0], [c_1]) = \inf_{\phi \in \mathcal{D}} d_{\mathcal{M}}(c_0, c_1 \circ \phi) \tag{28}$$

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. Some precautions has to be taken here: whereas \mathcal{M} is a submanifold of the Fréchet manifold $\mathcal{C}^\infty([0, 1], SO(n))$, has proven by [52], Theorem 10.4, the shape space \mathcal{S} is not a manifold and the principal bundle structure $\pi = \mathcal{M} \rightarrow \mathcal{S}$ is not formally defined. However, a manifold structure can be obtained if we only consider free immersion [53]. As the metric defined on the shape space is reparametrization-invariant, it is constant along the “fibers” (the origin point is fixed). Further explanations on the Riemannian submersion can be found in [54]. Closed curves being of main interest in this work, we can also define the set

$$\mathcal{M}^c = \{c \in \mathcal{C}([0, 1], SO(n)) : c'(t) \neq 0, c(0) = c(1)\}. \tag{29}$$

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

$$\mathcal{M}^{c+} = \{c \in \mathcal{C}([0, 1], SO(n)) : c'(t) \neq 0, c(0) = c(1), c'(0) = c'(1)\}. \tag{30}$$

We need now to introduce the Square Root Velocity function (SRV function) [55], in which a curve is represented by its starting point and its normalized velocity at each time t . There are several possibilities to define the SRV of a curve. The more general definition is the following

$$F : \mathcal{M} \rightarrow SO(n) \times T\mathcal{M} \\ c \rightarrow \left(c(0), q = \frac{c'}{\sqrt{\|c'\|}} \right). \tag{31}$$

However, we can go further and benefit from the specific case of the Lie group. In this section, we will denote the base manifold $G = SO(n)$ to emphasize its group structure, and g an element of the group. As in [47], 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 [47,48,56] 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 [47], a TSRV function of the following form:

$$F_{Lie} : C^\infty([0, 1], G) \longrightarrow SO(n) \times \{q \in C^\infty([0, 1], \mathfrak{g}), q(t) \neq 0, \forall t \in [0, 1]\}$$

$$F_{Lie}(c)(t) = (c(0), q(t)) = \left(c(0), \frac{R_{c(t)^*}^{-1}(c'(t))}{\sqrt{\|c'(t)\|}} \right) = \left(c(0), \frac{T_c^{c(t) \rightarrow I}(c'(t))}{\sqrt{\|c'(t)\|}} \right), \tag{32}$$

where \mathfrak{g} is the Lie algebra, R is the right translation on the group, $R_{g_1}(g_2) = g_2g_1$, $R_{g^*} = T_eR_g$ is the tangent map at the identity, $\|\cdot\|$ is a norm induced by a right-invariant metric on G , and $T_c^{c(t) \rightarrow I}$ denotes the parallel transport from $c(t)$ to the identity according to the curve c . A curve is now represented pointwise as an element of the tangent bundle $(c(0), q(t)) \in M \times \mathfrak{g}$ (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 $q \in C^\infty([0, 1], T\mathcal{M})$, 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 $SO(n)$.

3.3. 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 following development and expression of metrics and distances can be found in [45]. The distance on the shape space is used to compare how the curves are intrinsically different. It has been seen in [57] that the simple L^2 metric on \mathcal{M} given by

$$g_c^{L^2}(u, v) = \int \langle u, v \rangle \|c'(t)\| dt \tag{33}$$

where $\langle \cdot, \cdot \rangle$ is the Riemannian metric on $SO(n)$, induced a vanishing metric on the shape space, that is, we cannot differentiate shape with this metric. To overcome this difficulty, the family of *elastic metric*, derived from the Sobolev metric [58,59], has been investigated for it is non-vanishing on the shape space. In the case of a Euclidean space \mathbb{R}^n , it admits the expression:

$$g_c^{a,b}(u, v) = \int \left(a^2 \langle D_1u^N, D_1v^N \rangle + b^2 \langle D_1u^T, D_1v^T \rangle \right) \|c'(t)\| dt, \tag{34}$$

where $D_1u = h' / \|c'\|$, $D_1u^T = \langle D_1u, w \rangle w$, with $w = c' / \|c'\|$ and $D_1u^N = D_1u - D_1u^T$. Here, we are only interested in the special metric that has been proposed in [45], and which is an adaptation of the *elastic metric* for the Riemannian manifold. In our case this gives:

$$g_c(u, v) = \int \left(\langle \nabla_1u^N, \nabla_1v^N \rangle + \frac{1}{4} \langle \nabla_1u^T, \nabla_1v^T \rangle \right) \|c'(t)\| dt, \tag{35}$$

where ∇ is the Levi-Civita connection that corresponds to $\langle \cdot, \cdot \rangle$; $\nabla_1u = \frac{1}{\|c'\|} \nabla_c h$, $\nabla_1u^T = \langle \nabla_1u, w \rangle w$, $w = c' / \|c'\|$. For the computations being done now in a manifold space, the Levi-Civita connection replaced the ordinary derivative of \mathbb{R}^n .

Once geometry has been settled in \mathcal{M} , the geometry of the shape space can be derived from its quotient structure. Let the tangent bundle be decomposed into a vertical and a horizontal subspace: $T\mathcal{M} = \mathcal{H}_\mathcal{M} \oplus \mathcal{V}_\mathcal{M}$, with $\mathcal{V}_\mathcal{M} = \ker(T_c\pi)$ and T_c the tangent map, $\pi : \mathcal{M} \rightarrow \mathcal{S}$ the principal bundle, and $\mathcal{H}_\mathcal{M} = (\mathcal{V}_\mathcal{M})^\perp$, see Figure 3. This metric is reparametrisation invariant, that is, constant along the fibers, hence we have

$$g_c(u_\mathcal{H}, v_\mathcal{H}) = [g]_{\pi(c)}(T_c\pi(u), T_c\pi(v)) \tag{36}$$

where $[g]$ denotes the metric on the shape space.

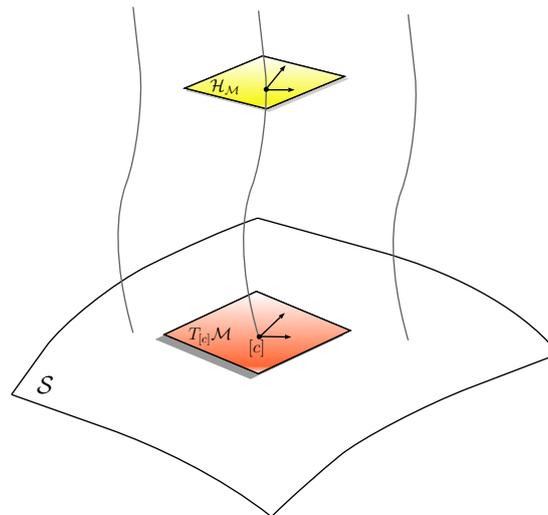


Figure 3. The tangent space $T_{[c]}\mathcal{M}$ at a point $[c]$ in the shape space \mathcal{S} is isomorphic to the horizontal part \mathcal{H}_M of the tangent space at a point on the associated fiber.

A similar result in a different (but still close) context is used in Lemma 1 of [60]. 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 reparametrized version of c_1 , $c_1 \circ \phi$ which gives the distance in the shape space:

$$[d]([c_0], [c_1]) = d_g(c_0, c_1 \circ \phi) \tag{37}$$

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 Equation (37) yields an optimisation problem:

$$[d]([c_0], [c_1]) = \inf_{\phi \in \mathcal{D}} \left(\int_0^1 \| q_0(t) - q_1(\phi(t)) \sqrt{\phi'(t)} \|^2 \right)^{1/2}, \tag{38}$$

which is solved by a traditional gradient descent algorithm or a dynamic linear programming [47]. Finally, we have to mention that in a practical situation, the above formula has to be discretized. This is the object of [46]. Formulae are essentially similar, but in this setting, a curve is now represented by a set of points $c_{disc}(x_0, x_1, \dots, x_n)$ and the tangent space turns into

$$T_{disc}\mathcal{M} = \{v = (v_0, v_1, \dots, v_n), v_i \in T_{x_i}SO(n), \forall i\}. \tag{39}$$

Concerning the metric on the space of curves, it becomes

$$g_{c_{disc}}(u, v) = \langle u_0, v_0 \rangle + \frac{1}{n} \sum_{i=0}^{n-1} \langle \nabla_{\partial c / \partial s} q^u \left(0, \frac{k}{n} \right), \nabla_{\partial c / \partial s} q^v \left(0, \frac{k}{n} \right) \rangle \quad \forall u, v \in T_{disc}\mathcal{M} \tag{40}$$

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 discretized 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

$$\nabla_{\partial c / \partial s} q \left(s, \frac{k}{n} \right) = \nabla_{\partial c / \partial s} q_k(s) \tag{41}$$

3.4. The Geodesic Equation

Let us now 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 [45,46]. The formula can be found in [47]. For the sake of completeness, we give a reformulated proof in Appendix B. Recall that a geodesic is a particular path of curves. A path of curves 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 A1). 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): $q = \frac{T_c^{c(t) \rightarrow I}(\partial c / \partial t)}{\sqrt{\|\partial c / \partial t\|}}$

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

$$\nabla_{\partial c / \partial s} (\nabla_{\partial c / \partial s} q(s, t)) (s, t) = 0 \quad \forall s, t \tag{42}$$

Proof. Appendix B \square

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

$$F_{Lie}^{-1} ((1 - s)F_{Lie}(c_0) + sF_{Lie}(c_1)) \tag{43}$$

for $s \in [0, 1]$. This expression is nothing but a linear interpolation on the transported tangent space.

We now have all the ingredients to give the procedure for nonstationary processes characterization and comparison:

1. **Input:** a set of rotation matrices $\{W_i\}_i$, seen as a partial observation of a closed trajectory on $SO(n)$.
2. Map the set of matrices $\{W_i\}$ into the a set of matrices in the Lie algebra $\{V_i\}$ using the inverse exponential map.
3. Interpolate with splines between matrices V_i [61,62].
4. Go back in the base manifold $SO(n)$ with the exponential map.
5. Shift the interpolated curve in order to fulfill the condition $c(0) = e$ and compute the SRV transformation given by (41)
6. Compute the distance defined by (38). The optimization is carried out by dynamic programming.
7. **Output:** distance between two curves in the manifold defined by the set of curves in $SO(n)$, and geodesic path between the curves.

The interpolation computations are carried out in the Lie algebra, which is a vector space, and thus it does not demand great computational resources. The discretization step, which amounts to choosing certain values among the continuous curve, is also done in the Lie algebra. In this way, we avoid the calculation of matrices exponential that would have been discarded at the end. We finally note that geodesic shooting [45,63] or other path straightening methods could be applied to obtain a geodesic path between two curves, and between the shapes of the two curves.

3.5. 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 4. We also bring their corresponding $SO(3)$ representation on Figures 5 and 6 with 200 interpolated points and 50 interpolated points respectively. 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) (AutoRegressive) with a period of 54 points, and a periodic ARMA(2,1) (AutoRegressive Moving Average) 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 three parcors each. The analysis of Figure 4 with Figure 5 shows that the spectral measure of the amplitude modulated signal of Figure 4a 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 4d is quite identical to the amplitude modulated signal of Figure 4a, 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 4b,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. We also note one of the advantages of using spline interpolation. Along with Figure 5, we can remark that the curvature is well approximated owing to more and closer interpolated points. For this figure, we had approximately four times more interpolated points than original ones whereas for Figure 6 we computed roughly twice more interpolated points than the original ones. Actually, for such a number, a change in the grid could be thought but it seems again that the spline interpolation gives good points repartition. This is mainly illustrated by Figure 7 for which the difference between geodesic interpolation of the curve associated with the periodic AR(2) signal of Figure 4c for 200 interpolated points and 50 interpolated points is very weak.

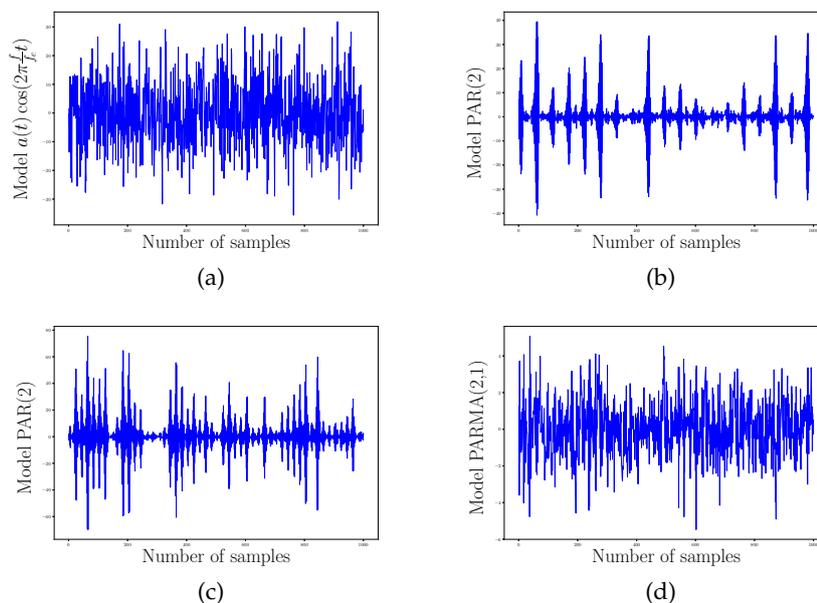


Figure 4. 1000 samples of PC processes generated by (a) a modulated zero mean and unit variance stationary random process $a(t)$; (b) a periodic AR(2) model with a period of 54 points; (c) a periodic AR(2) model with a period of 20 points; and (d) a periodic ARMA(2,1) model with a period of 20 points.

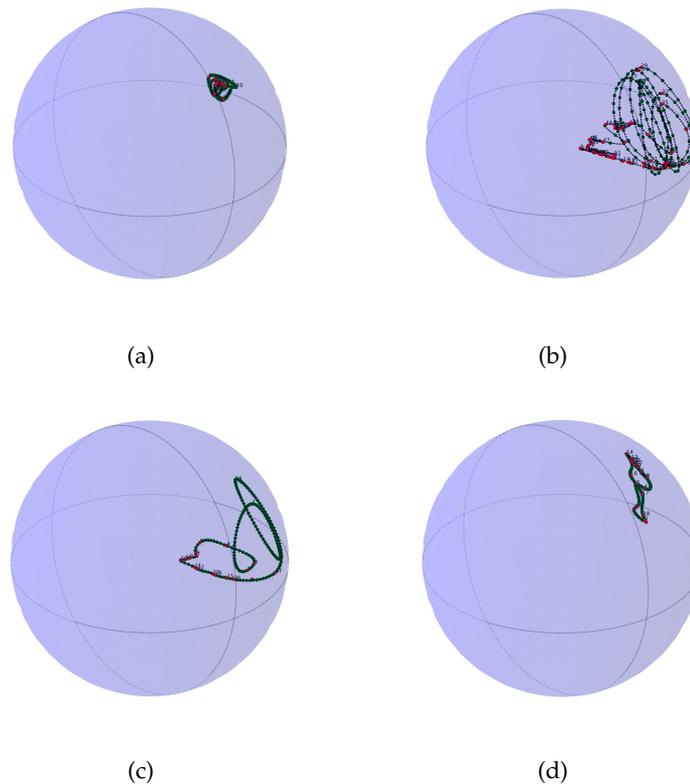


Figure 5. Representation inside the ball of radius π of the four PC processes drawn in Figure 4, arranged in the same order with 200 interpolated points represented with green stars, the dashed black line is the theoretical curve and the red dots are the representation of dilation matrices. (a) is the representation of Figure 4a, (b) is the representation of Figure 4b, (c) is the representation of Figure 4c and (d) is the representation of Figure 4d.

To end this analysis by the example, we have computed the distance defined by (38) between the PC process of Figure 8 and all the PC processes studied and displayed on Figures 4 and 5. The distances are reported inside 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. We have gray colored the row of the PAR(2) signal model indexed by letter (c). Whatever the number of interpolated points and the dimension of the base manifold, the spectral representation through dilation operators of this signal is the nearest on $SO(3)$ and the second nearest for $SO(4)$ and $SO(5)$ to the PAR(2) signal of reference. The second interesting signal is the one indexed by letter (b) and stands for a PAR(2) signal with exactly the same model parameters as that of the signal of reference but with 54 points of periodicity instead of 20. We have lightly gray colored its associated row on Table 1 when it had the shortest distance. The curve associated with this signal has many wraps on its representation, and it has consequently the greatest distance on $SO(3)$, but increasing the dimension improves the comparison. Indeed, it finally has the shortest distance on $SO(4)$ and $SO(5)$. This is particularly interesting to see that there is a competition between the curves of a PAR(2) with different model parameters but the same periodicity and a PAR(2) with the same model parameters but different periodicity.

Table 1. Table of the distances between all the PC processes of Figure 4 to the gold standard PC process of Figure 8 through the distance of their curves' shapes on $SO(3)$, $SO(4)$ and $SO(5)$. We have interpolated with roughly twice and four times the number of original points. We also applied here a DP to solve the optimization assignment problem.

Model of Signal Displayed in Figure 4	Distance to the Signal of Figure 8					
	SO(3)		SO(4)		SO(5)	
	200 pts	50 pts	200 pts	50 pts	200 pts	50 pts
(a)	5.72	4.47	97.19	26.86	526.36	95.47
(b)—100 pts instead of 50 pts	31.63	28.98	41.78	12.98	298.64	220.32
(c)	3.44	3.29	90.89	20.23	476.55	116.06
(d)	4.19	4.50	187.42	50.36	621.51	171.73

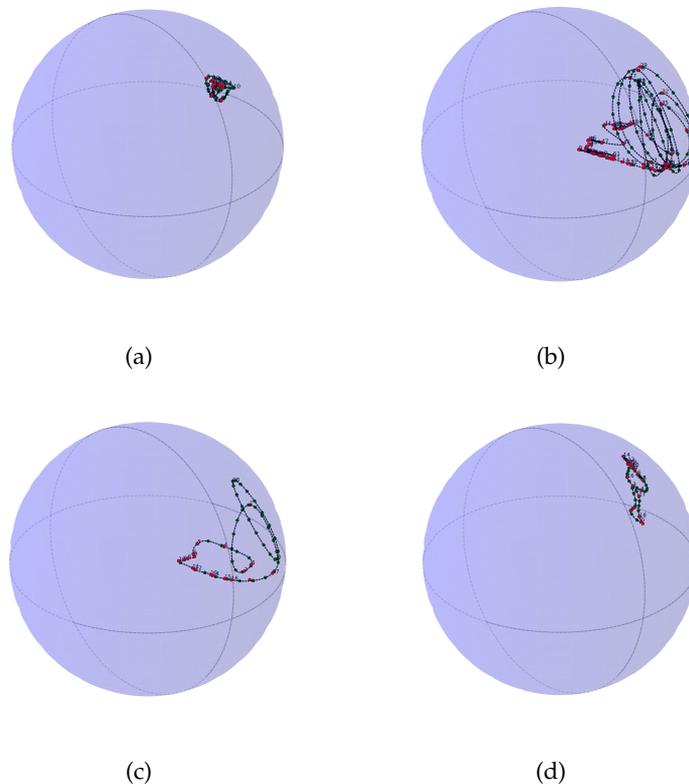


Figure 6. Representation inside the ball of radius π of the four PC processes drawn in Figure 4, arranged in the same order with 50 interpolated points except for plot (b) which has been computed with 100 interpolated points represented with green stars; the dashed black line is the theoretical curve and the red dots are the representation of dilation matrices. (a) is the representation of Figure 4a, (b) is the representation of Figure 4b, (c) is the representation of Figure 4c and (d) is the representation of Figure 4d.

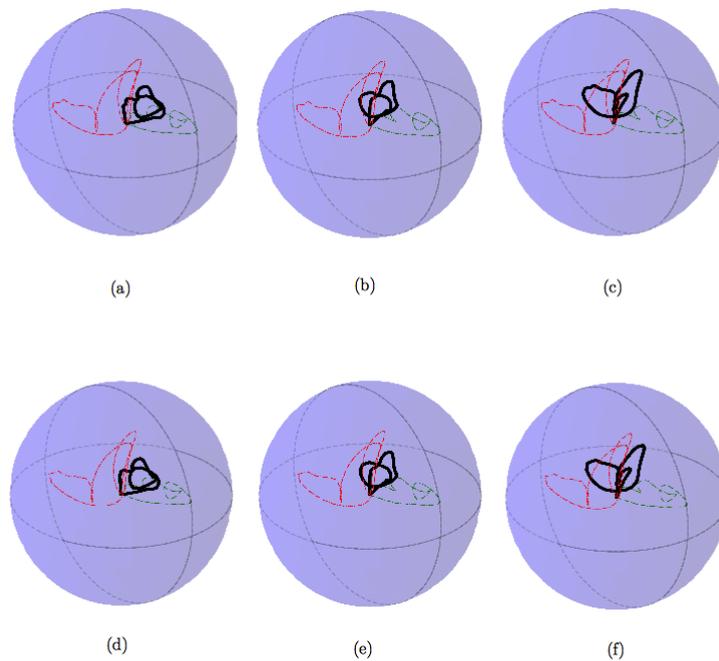


Figure 7. Geodesic interpolation with respect to (43) between the green dashed curve (gold standard signal of Figure 8) and the dashed red curve (signal of Figure 4c), first row for 200 interpolated points, second row for 50 interpolated points. For this scenario $s \in [1/4, 1/2, 3/4]$ which corresponds to [(a–d), (b–e), (c–f)] respectively.

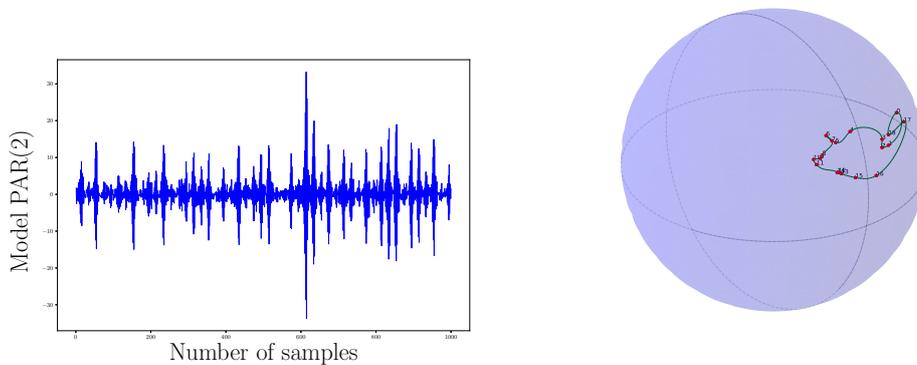


Figure 8. A PAR(2) signal with a period of 20 points, 1000 samples were generated, and its corresponding $SO(3)$ representation inside the ball of radius π .

We end by noticing that the PARMA(2,1) has the second shortest distance to the PAR(2) signal model of reference on $SO(3)$. As Figure 5 shows, their spectral measure evolves in a similar way with one major loop and a second less important loop. However, once the dimension of the base manifold increases, the assumption that the two processes may be close is strongly rejected by the fact that the PARMA(2,1) signal model has the longest distance. Finally, these observations leave open besides the question of the topology of these curves and how it could be used for the classification.

4. Conclusions

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 angles correspond to the well-known parcos, 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 $SO(n)$ 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, M.D.; Formal analysis, M.D. and G.B.; Funding acquisition, G.B. and E.M.; Investigation, M.D.; Methodology, G.B.; Supervision, G.B. and E.M.; Writing—original draft, M.D.; Writing—review and editing, G.B.

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:

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

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 A1. *Let X and Y be operators in \mathcal{z} . The following statements are equivalent:*

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

Proof. This result can be proved by taking the contraction Γ with respect to $\Gamma Xh = Yh$. [41]. \square

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 intervenes 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 start with a basic case:

Theorem A2 (row contraction). *Let $T = [T_1 \ T_2] \in \mathcal{L}(\mathcal{H}_1 \oplus \mathcal{H}_2, \mathcal{H})$, then $\|T\| \leq 0$ 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*

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

Proof. The proof is a simple application of Theorem A1. For the if part, it is obvious that we can take Γ_1 to be T_1 . Then $\|T\| \leq 1$ implies

$$I - TT^* = I - \Gamma_1 \Gamma_1^* - T_2 T_2^* \geq 0 \tag{A3}$$

with $D_{\Gamma_1^*}^2 \geq T_2 T_2^*$. Hence, there exists Δ such that $\Delta D_{\Gamma_1^*} = T_2^*$. Choosing $\Gamma_2 = \Delta^*$ finishes the argument. \square

In the same way as that of the Cholesky factorisation, we can write down the defect operator for the whole contraction $T = [T_1 \ T_2]$ [41] 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 A2, 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,

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

and the general case is:

Theorem A3 (Structure of row contraction). *The following are equivalent:*

- The operator $T^n = [T_1 \ T_2 \ \dots \ T_n]$ 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 $T_k = D_{\Gamma_1^*} D_{\Gamma_2^*} \dots D_{\Gamma_{k-1}^*} \Gamma_k$.

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

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

and

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

Proof. It can be proved straightforwardly by induction. \square

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 A4. 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:

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

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

Similarly to [45,48], 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 [45]. Here, we work directly in the tangent space representation, via the SRV representation, and not with “the whole family” of curves $c(s, t, \tau)$.

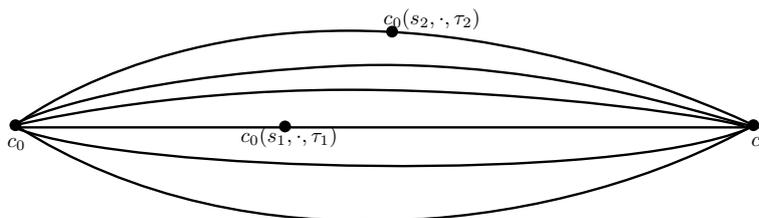


Figure A1. we consider a beam of curves, which consists in a slight modification of the geodesic. The different curves are indexed by τ . The idea is to find which of these curves gives the minimal energy to go from c_0 to c_1 .

We denote $\partial_{\tau} c(s, 0, \tau) = \frac{\partial c(s, 0, \tau)}{\partial \tau}$, 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

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

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

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

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

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

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

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

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. Auger, F.; Flandrin, P.; Lin, Y.-T.; McLaughlin, S.; Meignen, S.; Oberlin, T.; Wu, H.-T. Time-Frequency Reassignment and Synchrosqueezing: An Overview. *IEEE Signal Process. Mag.* **2013**, *30*, 32–41. [[CrossRef](#)]
2. Flandrin, P. *Time-Frequency/Time-Scale Analysis*; Academic Press: Cambridge, MA, USA, 1999; Volume 10.
3. Hurd, H.L.; Miamee, A. *Periodically Correlated Random Sequences: Spectral Theory and Practice*; Wiley Series in Probability and Statistics; John Wiley: Hoboken, NJ, USA, 2007.
4. Napolitano, A. Cyclostationarity: New trends and applications. *Signal Process.* **2016**, *120*, 385–408. [[CrossRef](#)]
5. Dégerine, S.; Lambert-Lacroix, S. Characterization of the partial autocorrelation function of nonstationary time series. *J. Multivar. Anal.* **2003**, *87*, 46–59. [[CrossRef](#)]
6. Lambert-Lacroix, S. Extension of Autocovariance Coefficients Sequence for Periodically Correlated Processes. *J. Time Ser. Anal.* **2005**, *26*, 423–435. [[CrossRef](#)]
7. Desbouvries, F. Unitary Hessenberg and state-space model based methods for the harmonic retrieval problem. *IEE Proc. Radar Sonar Navig.* **1996**, *143*, 346–348. [[CrossRef](#)]
8. Yang, L.; Arnaudon, M.; Barbaresco, F. Riemannian median, geometry of covariance matrices and radar target detection. In Proceedings of the 7th European Radar Conference, Paris, France, 30 September–1 October 2010; pp. 415–418.
9. Bingham, N.H. Szegő's Theorem and Its Probabilistic Descendants. *Prob. Surv.* **2012**, *9*, 287–324. [[CrossRef](#)]
10. Simon, B. *Orthogonal Polynomials on the Unit Circle Part 1 and Part 2*; American Mathematical Society: Providence, RI, USA, 2009; Volume 54.
11. Delsarte, P.; Genin, Y.V.; Kamp, Y.G. Orthogonal polynomial matrices on the unit circle. *IEEE Trans. Circuits Syst.* **1978**, *25*, 149–160. [[CrossRef](#)]

12. Barbaresco, F. Radar micro-Doppler Signal Encoding in Siegel Unit Poly-Disk for Machine Learning in Fisher Metric Space. In Proceedings of the 2018 19th International Radar Symposium (IRS), Bonn, Germany, 20–22 June 2018.
13. Simon, B. CMV matrices: Five years after. *J. Comput. Appl. Math.* **2007**, *208*, 120–154. [[CrossRef](#)]
14. Ammar, G.; Gragg, W.; Reichel, L. *Constructing a Unitary Hessenberg Matrix from Spectral Data*; Springer: Berlin, Germany, 1991; pp. 385–395.
15. Masani, P. Dilations as Propagators of Hilbertian Varieties. *SIAM J. Math. Anal.* **1978**, *9*, 414–456. [[CrossRef](#)]
16. Constantinescu, T. *Schur Parameters, Factorization and Dilation Problems*; Birkhäuser: Basel, Switzerland, 1995.
17. Nagy, B.S.; Foias, C.; Bercovici, H.; Kérchy, L. *Harmonic Analysis of Operators on Hilbert Space*; Springer: New York, NY, USA, 2010.
18. Arsene, G.; Constantinescu, T. The Structure of the Naimark Dilation and Gaussian Stationary Processes. *Integral Equ. Oper. Theory* **1985**, *8*, 181–204. [[CrossRef](#)]
19. Arnaudon, M.; Barbaresco, F.; Yang, L. Riemannian Medians and Means With Applications to Radar Signal Processing. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 595–604. [[CrossRef](#)]
20. Barbaresco, F. 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: Berlin, Germany, 2009; pp. 124–163.
21. Desbouvries, F. Geometrical Aspects of Linear Prediction Algorithms. Available online: <http://www.math.ucsd.edu/helton/MTNSHISTORY/CONTENTS/2000PERPIGNAN/CDROM/articles/B85.pdf> (accessed on 9 September 2018)
22. Desbouvries, F. Non-euclidean geometrical aspects of the schur and levinson-szego algorithms. *IEEE Trans. Inf. Theory* **2003**, *49*, 1992–2003. [[CrossRef](#)]
23. Balaji, B.; Barbaresco, F.; Decurninge, A. Information geometry and estimation of Toeplitz covariance matrices. In Proceedings of the 2014 International Radar Conference, Lille, France, 13–17 October 2014; pp. 1–4.
24. Barbaresco, F. Geometric Radar Processing based on Fréchet distance: Information geometry versus Optimal Transport Theory. In Proceedings of the 2011 12th International Radar Symposium (IRS), Leipzig, Germany, 7–9 September 2011; pp. 663–668.
25. Barbaresco, F. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median. In *Matrix Information Geometry*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 199–255.
26. Lipeng, N.; Xianhua, J.; Georgiou, T. Geometric methods for structured covariance estimation. In Proceedings of the 2012 American Control Conference (ACC), Montreal, QC, Canada, 27–29 June 2012; pp. 1877–1882.
27. Pennec, X. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *J. Math. Imaging Vis.* **2006**, *25*, 127–154. [[CrossRef](#)]
28. Pennec, X.; Fillard, P.; Ayache, N. A Riemannian framework for tensor computing. *Int. J. Comput. Vis.* **2006**, *66*, 41–66. [[CrossRef](#)]
29. Barbaresco, F.; Ruiz, M. Radar detection for nonstationary Doppler signal in one burst based on information geometry: Distance between paths on covariance matrices manifold. In Proceedings of the 2015 European Radar Conference (EuRAD), Paris, France, 9–11 September 2015; pp. 41–44.
30. Le Brigant, A.; Arnaudon, M.; Barbaresco, F. Optimal Matching Between Curves in a Manifold. In Proceedings of the 2017 International Conference on Geometric Science of Information, Paris, France, 7–9 November 2017; pp. 57–64.
31. Arnold, V. Sur la Géométrie Différentielle des Groupes de Lie de Dimension Infinie et Ses Applications à L'hydrodynamique des Fluides Parfaits. *Annales de l'institut Fourier* **1966**, 319–361. (In French) [[CrossRef](#)]
32. Kailath, T.; Bruckstein, A.M. Naimark Dilations, State-Space Generators and Transmission Lines. In *Advances in Invariant Subspaces and Other Results of Operator Theory*; Douglas, R.G., Percy, C.M., Nagy, B.S., Vasilescu, F.-H., Voiculescu, D., Arsene, G., Eds.; Birkhäuser: Basel, Switzerland, 1986; pp. 173–186.
33. Sayed, A.H.; Constantinescu, T.; Kailath, T. Recursive Construction of Multichannel Transmission Lines with a Maximum Entropy Property. In *Codes, Graphs, and Systems*; Springer: Berlin, Germany, 2002; pp. 259–290.
34. Makagon, A.; Salehi, H. Notes on infinite dimensional stationary sequences. In *Probability Theory on Vector Spaces IV*; Cambanis, S., Weron, A., Eds.; Lecture Notes in Mathematics; Springer: Berlin/Heidelberg, Germany, 1989; Volume 1391, pp. 200–238.

35. Miamee, A. Spectral dilation of $L(B,H)$ -valued measures and its application to stationary dilation for Banach space valued processes. *Indiana Univ. Math. J.* **1989**, *38*, 841–860. [[CrossRef](#)]
36. Miamee, A. Periodically Correlated Processes and Their Stationary Dilations. *SIAM J. Appl. Math.* **1990**, *50*, 1194–1199. [[CrossRef](#)]
37. Bochner, S. *Vorlesungen über Fouriersche Integrale*; Mathematik und ihre Anwendungen, Akademie-Verlag: Leipzig, Germany, 1932.
38. Foias, C.; Frazho, A.E. A Geometric Approach to Positive Definite Sequences. In *The Commutant Lifting Approach to Interpolation Problems*; Birkhäuser: Basel, Switzerland, 1990; pp. 497–546.
39. Frazho, A.E.; Arthur, E. On Stochastic Bilinear Systems. In *Modelling and Application of Stochastic Processes*; Springer: Boston, MA, USA, 1986; pp. 214–241.
40. Timotin, D. Prediction theory and choice sequences: An alternate approach. In *Advances in Invariant Subspaces and Other Results of Operator Theory*; Springer: Berlin, Germany, 1986; pp. 341–352.
41. Tseng, M.C. Contractions, Matrix Parametrizations, and Quantum Information. *arXiv* **2006**, arXiv:quant-ph/0610259.
42. Tseng, M.C.; Ramakrishna, V. Dilation Theoretic Parametrizations of Positive Matrices with Applications to Quantum Information. *arXiv* **2006**, arXiv:quant-ph/0610021.
43. Damanik, D.; Pushnitski, A.; Simon, B. The Analytic Theory of Matrix Orthogonal Polynomials. *Surv. Approx. Theory* **2008**, *4*, 1–85.
44. Bakonyi, M.; Constantinescu, T. Pitman Research Notes in Math. In *Schur's Algorithm and Several Applications*; Longman Sc and Tech: Harlow, UK, 1992; Volume 261.
45. Le Brigan, A. Computing distances and geodesics between manifold-valued curves in the SRV framework. *arXiv* **2016**, arXiv:1601.02358.
46. Le Brigan, A. A discrete framework to find the optimal matching between manifold-valued curves. *arXiv* **2017**, arXiv:1703.05107.
47. Celledoni, E.; Eslitzbichler, M.; Schmeding, A. Shape Analysis on Lie Groups with Applications in Computer Animation. *arXiv* **2015**, arXiv:1506.00783.
48. Zhang, Z.; Su, J.; Klassen, E.; Le, H.; Srivastava, A. Video-based action recognition using rate-invariant analysis of covariance trajectories. *arXiv* **2015**, arXiv:1503.06699.
49. Le Brigan, A.; Arnaudon, M.; Barbaresco, F. A Probability on the Spaces of Curves and the Associated Metric Spaces via Information Geometry; Radar Application. Ph.D. Thesis, Université de Bordeaux, Bordeaux, France, 4 July 2017.
50. Kobayashi, S.; Nomizu, K. *Foundations of Differential Geometry*; Interscience Publishers: New York, NY, USA, 1963; Volume 1.
51. Kriegl, A.; Michor, P.W. Aspects of the theory of infinite dimensional manifolds. *Differ. Geom. Appl.* **1991**, *1*, 159–176. [[CrossRef](#)]
52. Michor, P.W.; Mumford, D. Shiva Mathematics Series. In *Manifolds on Differential Mappings*; Birkhauser: Basel, Switzerland, 1980; Volume 3.
53. Michor, P.W.; Mumford, D. An overview of the Riemannian metrics on shape spaces of curves using the Hamiltonian approach. *Appl. Comput. Harmon. Anal.* **2007**, *23*, 74–113. [[CrossRef](#)]
54. Michor, P.W. *Topics in Differential Geometry*; Graduate Studies in Mathematics American Mathematical Society: Providence, RI, USA, 2008; Volume 93.
55. Srivastava, A.; Klassen, E.; Joshi, S.H.; Jermyn, I.H. Shape Analysis of Elastic Curves in Euclidean Spaces. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 1415–1428. [[CrossRef](#)] [[PubMed](#)]
56. Bauer, M.; Bruveris, M.; Marsland, S.; Michor, P.W. Constructing reparameterization invariant metrics on spaces of plane curves. *Differ. Geom. Appl.* **2014**, *34*, 139–165. [[CrossRef](#)]
57. Michor, P.W.; Mumford, D. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math* **2005**, *10*, 217–245.
58. Bauer, M.; Bruveris, M.; Michor, P.W. Why Use Sobolev Metrics on the Space of Curves. In *Riemannian Computing in Computer Vision*; Springer: Cham, Switzerland, 2016; pp. 233–255.
59. Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*; CRC Press: Boca Raton, FL, USA, 2014; Volume 2.
60. Yili, L.; Wong, K.M. Riemannian Distances for Signal Classification by Power Spectral Density. *IEEE J. Sel. Top. Signal Process.* **2013**, *7*, 655–669. [[CrossRef](#)]

61. Hofer, M.; Pottmann, H. Energy-minimizing Splines in Manifolds. *ACM Trans. Graph.* **2004**, *23*, 284–293. [[CrossRef](#)]
62. Shingel, T. Interpolation in special orthogonal groups. *IMA J. Num. Anal.* **2009**, *29*, 731–745. [[CrossRef](#)]
63. Pilté, M.; Barbaresco, F. Tracking quality monitoring based on information geometry and geodesic shooting. In Proceedings of the 2016 17th International Radar Symposium (IRS), Krakow, Poland, 10–12 May 2016; pp. 1–6.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).