

Representation and Characterization of Non-Stationary Processes by Dilation Operators and Induced Shape Space Manifolds

Maël Dugast, Guillaume Bouleux, Eric Marcon

Abstract

We have introduced a new vision of stochastic processes through the geometry induced by the dilation. The dilation matrices of a given process which are obtained by a composition of rotation matrices, contain the measure information in a condensed way. 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, it coincides with the Naimark Dilation. When the process is non-stationary, it returns a set of rotation matrices. In particular, the periodicity of the correlation function that may appear in some classes of signal is transmitted to the set of dilation matrices. These rotation matrices, which can be arbitrarily close to each other depending on the sampling or the rescaling of the signal are seen as a distinctive feature of the signal. In order to study this sequence of matrices, and guided by the possibility to rescale the signal, the correct geometrical framework to use with the dilation's theoretic results is the space of curves on manifolds, that is the set of all curves that lie on a based manifold. To give a complete sight about the space of curves, a metric and the derived geodesic equation are provided. The general results are adapted to the more specific case where the base manifold is the Mie group of rotation matrices. The notion of the shape of a curve can be formalized as the set of equivalence classes of curves given by the quotient space of the space of curves and the increasing diffeomorphisms. The metric in the space of curves naturally extends to the space of shapes and enables comparison between shapes.

1 Introduction

The Analysis and/or the representation of non-stationary processes has been tackled during 4 or 5 decades for now by time-scale/time frequency analysis [21, 4], by Fourier-like representation when the processes belong to the periodically-correlated subclass [39, 24] or by partial correlation coefficients (pacors) series [29, 20] to cite a few. One of the advantages of dealing with pacors resides in their strong relation to the measure of the process by the one-to-one relation with correlation coefficients [18, 52]. They consequently appear explicitly in the Orthogonal Polynomial on the Real Line (OPRL)/Unit Circle (OPUC) decomposition of the measure [11, 44] and are the elements for the construction of dilation matrices that appear in the CMV/GGT [43], for the Schur flows problem with upper Hessenberg matrices [1] which are also seen in the literature as evolution operators [44] or shift operator [34], and appear finally in the state-space representation [15, 17]. The dilation theory takes its roots in the Operators theory [48] which bridges the process's measure and unitary operators. In its simplest version, the dilation theory corresponds to the Naimark dilation [48, 3], 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 [15] 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 non-stationary case, by a sampled trajectory induced by the dilation matrices on the appropriate Lie Group. When the process is periodically correlated, the sequence of pacors inherits of the periodicity and the sequence of dilation matrices becomes periodic as well, we obtain consequently a closed path as illustrated in Fig.1. Characterizing 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 for describing stochastic processes [33]. The second order statistical properties/moments may be analyzed, characterized and compared [8, 5] to improve estimation [47, 38] or classification of different processes [27]. When dealing with density estimation [25], the space

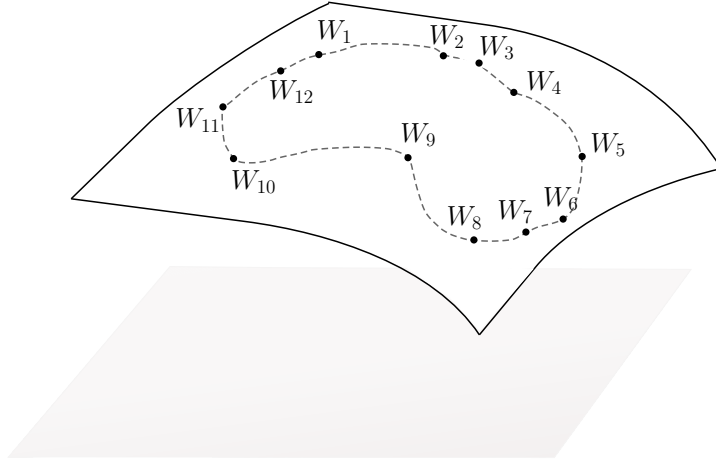


Figure 1: Illustration of a sampled closed trajectory drawn in $SO(n)$ or $SU(n)$ which materializes the time varying of the periodically-correlated measure for a stochastic process. Each W_i are dilation matrix built through the parcors.

of $n \times n$ symmetric matrices $Sym(n)$ is generally preferred and many developments have been proposed under the Semi Positive Definite (SPD) assumption [40, 41, 45, 14] for which the set of SPD matrices constitute a convex half-cone in the vector space of matrices. This leads to give more insights on the Fisher Information metric [14, 25] or Wasserstein metric [28] and also cope with optimal mass transportation problems [7]. Many efforts have also been made 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 [19, 16, 52, 6, 2]. 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 non-stationary 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 characterization of processes’ measure in non-stationary context. First, using the dilation theory approach to give sampled points and next giving the prescribed geodesics equations used for curves or paths comparisons in 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 [26, 42] are also discussed and the strong connection between parcors and the dilation matrices are emphasized. The 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 Lie Group and a procedure to compare non-stationary 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 on the dilation theory. In its fundamental definition, the dilation theory consists for an Hilbert space \mathcal{H} and an operator-valued function f , *i.e.* a $\mathcal{L}(\mathcal{H})$ -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} :

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

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 [32, 35, 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 result is that, given a contraction Γ , the aforementioned unitary Julia operator

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

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

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

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 [34] 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 a 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

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

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

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

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 [12, 3, 51]. Owing this property, we obtain the following relation

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

due to 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 know that 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 term of measure. If we denote by E_{λ} an operator-valued distribution function on $[0, 2\pi[$ then the function

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

This function is positive-definite and shows the strong correspondance between the spectral measure and the dilation theory. There exists so 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

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

or, in an equivalent form :

$$E_\lambda = P_{\mathcal{H}}F_\lambda. \quad (9)$$

Note that the operator-valued measure F_λ is in fact an orthogonal projection-valued measure since all its increments are orthogonal.

Dilations 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 a 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 $R_{i,j} = (1 \ 0 \ 0 \ \cdots \ 0) W_i W_{i+1} \cdots W_{j-1} (1 \ 0 \ 0 \ \cdots \ 0)^T$ 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
- There exists a unique contraction Γ in $\mathcal{L}(\mathcal{R}(Z), \mathcal{R}(X))$ such that

$$Y = X^{1/2} \Gamma Z^{1/2} \quad (10)$$

Proof. Annexe A □

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

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

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 $(R_{1,2} \ R_{1,3})$. 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 obtain thus a one-to-one correspondance 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 correspondance:

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

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

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

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

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

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

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

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 a 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

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

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

Where we have :

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

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

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

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

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

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

A different approach leading to the same results can be found in [49], using directly the Kolmogorov decomposition. In [26] the Naimark dilation is constructed using lattice filter and finally applications of this decomposition in quantum mechanics are to be found in [50, 51] for example.

2.3 Link between the Levinson algorithm and Constantinescu's formulae.

The theory explained so far provides infinite dimensional operators. Of course, it is not possible to treat with infinite dimensional object in practice, particularly with numerical computations. To strengthen our development, it is then necessary to adapt the theory to the finite dimensional case. This is tackled for example in [22] where a Levinson algorithm is performed on block matrices.

2.3.1 The principles of Levinson Algorithm

In classical stationary signal processing, the Levinson algorithm is a procedure to estimate linear regression coefficients, say $\{\alpha_k\}$ as the projection of the signal at time t onto the space spanned by its n past values: $X_t = -\sum_{i=1}^n \alpha_i X_{t-i} + \epsilon_t^n$, where ϵ is a white gaussian noise by assumption. The last coefficient of the model is the n -th reflexion (parcor) coefficient. If the correlation matrix of the stationary process is noted $R = \{R_{ij}\}_{i,j=1,\dots,n}$ we have the so called Yule-Waker equation:

$$\begin{pmatrix} R_0 & R_1 & \cdots & R_n \\ R_1^* & \ddots & \ddots & \\ \vdots & \ddots & \ddots & \\ R_n^* & \cdots & & R_0 \end{pmatrix} \begin{pmatrix} 1 \\ \alpha_1^n \\ \vdots \\ \alpha_n^n \end{pmatrix} = \begin{pmatrix} \epsilon_n^2 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \quad (17)$$

The Levinson algorithm performs a recursion on the order n of the model which appears to be exactly what is done for the dilation matrices.

2.3.2 Analogy between Levinson and Dilation

Let us recall that the form of the infinite dimensional dilation matrix is given by (15). This matrix acts on the space $\mathcal{K} = \mathcal{H} \oplus \mathcal{D}_1 \oplus \mathcal{D}_2 \cdots$ where the set of \mathcal{D}_i stands for the defect spaces. The forward prediction is defined by

$$U^n - P_{\mathcal{H} \oplus \mathcal{D}_1 \oplus \mathcal{D}_2 \cdots \oplus \mathcal{D}_n} U^n h = (0, 0, \dots, 0, \mathcal{D}_n \mathcal{D}_{n-1} \cdots \mathcal{D}_1 h, 0, \dots) \quad (18)$$

for $h \in \mathcal{H}$ and the non zero elements are in the $n+1$ position. In [22], the authors denotes the right-hand part by $\phi_n \epsilon_n h$ where $\epsilon_n = \mathcal{D}_n \mathcal{D}_{n-1} \cdots \mathcal{D}_1$ and ϕ_n embeds onto the $n+1$ position. They prove so the exact accordance with the classical Levinson procedure. In that case, $\epsilon_n = \mathcal{D}_n \mathcal{D}_{n-1} \cdots \mathcal{D}_1$ is therefore nothing else than the prediction error when the Levinson procedure is applied on the positive-definite kernel representing the correlation matrix.

Similarly, for the backward prediction error we have

$$h - P_{\cup C_n} h = \phi_n^* \epsilon_n^* h \quad (19)$$

where $\epsilon_n^* = \mathcal{D}_n^* \mathcal{D}_{n-1}^* \cdots \mathcal{D}_1^*$ and C_n stands for the n first columns of U . In terms of dilation, the Levinson procedure at step n can be written as

$$\epsilon_n^* \Gamma_{n+1} \epsilon_n h = R_{n+1} + \sum_{i=1}^n R_i A_{n,i-1} h \quad (20)$$

where the $A_{n,i}$ are the Levinson regression coefficients. Rewrite the equation such that

$$R_{n+1} = \epsilon_n^* \Gamma_{n+1} \epsilon_n h - \sum_{i=1}^n R_i A_{n,i-1} h \quad (21)$$

and compared it with theorem 2, allows to see that the Levinson recursion is equivalent to the parametrization of positive-definite kernel by means of Schur-Constantinescu parameters. We remark in addition that the necessary truncation error made when dealing with dilation matrices is now quantified.

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 exemple in [52] where the sequence of parcors asociated with a stationary process is seen as a point onto the Poincaré polydisk \mathcal{P}^n , *i.e.* the product of Poincaré disk. To give geometrical settings, a distance to characterize individual parcors is then proposed and discussed. In [30], 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 due to the product manifolds and the question of nonstationarity arises. Based on the work of [30, 31, 13] and [54], we propose then to give a particular attention to this question. We first make use of Dilation theory introduced in section 2. When the process under study is non-stationary, a set of matrices W_i is obtained. The basic idea for having geometric information on the non-stationary process is therefore to characterize the trajectory formed by the set of dilation matrices. These matrices are theoretically operator 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 which becomes perfect rotation operator belonging then to $SO(n)$ for real processes and $SU(n)$ when dealing with complex processes, when their dimension are reduced to $n \times n$. Our aim is finally to analyse those curves living on Lie Group of rotation matrices and emphasize the geometry or more precisely the intrinsic geometry formulation of these objects. For exemple, 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 of the computation complexity still rises 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 work of [30] and [13]. First we define the manifold \mathcal{M} given by the set of all curves in the based manifold. That leads to an other 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 followings.

3.1 Basic Outline of Geometry

Curves of interest are those ones living in the Lie group of real rotation matrices, this yields to $c : [0, 1] \rightarrow SO(n)$. For sake of clarity, we suppose that c is continuous, we will come back to the case of discrete curves later. To study geometrical features of such curves, we interest ourself in the set of all curves lying in $SO(n)$ (where $SO(n)$ is seen as a manifold) with non-vanishing 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 so a particular point in \mathcal{M} . The tangent space at a curve c is given by

$$T_c\mathcal{M} = \{v \in \mathcal{C}^\infty([0, 1], TSO(n)) : v(t) \in T_{c(t)}SO(n)\} \quad (22)$$

where $TSO(n)$ denote the tangent bundle of the based 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 this two curves should remain the same if the curve are only reparametrized, *i.e.* if we define other curves that passes 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 change the chosen points (see Figure 2). A reparametrization is represented by an 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

$$g_{c \circ \phi}(u \circ \phi, v \circ \phi) = g_c(u, v) \quad (23)$$

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

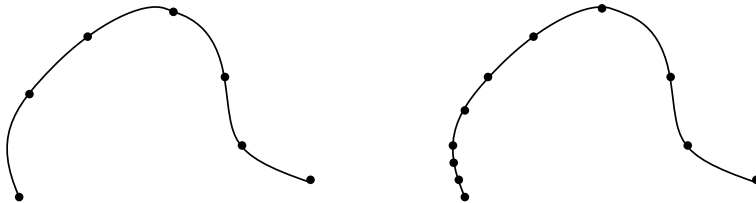


Figure 2: example of a reparametrization of a curve. Here, it consists in changing the discretization.

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

$$d_{\mathcal{M}}(c_0 \circ \phi, c_1 \circ \phi) = d_{\mathcal{M}}(c_0, c_1) \quad (24)$$

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} in such a way we have

$$c_0 \sim c_1 \iff \exists \phi \in \mathcal{D} : c_0 = c_1 \circ \phi. \quad (25)$$

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 writting

$$\mathcal{S} = \mathcal{M} / \sim, \text{ or } \mathcal{S} = \mathcal{M} / \mathcal{D}. \quad (26)$$

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

$$d_{\mathcal{S}}([c_0], [c_1]) = \inf_{\phi \in \mathcal{D}} d_{\mathcal{M}}(c_0, c_1 \circ \phi) \quad (27)$$

where $[c_0], [c_1]$ 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 $\pi = \mathcal{M} \rightarrow \mathcal{S}$. As closed curves are of main interest in this work, we can also define the set

$$\mathcal{M}^c = \{c \in \mathcal{C}([0, 1], SO(n)) : c'(t) \neq 0, c(0) = c(1)\}. \quad (28)$$

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

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

We need now to introduce the Square Root Velocity function (SRV function) [46], 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

$$F : \mathcal{M} \rightarrow SO(n) \times T\mathcal{M} \\ c \rightarrow \left(c(0), q = \frac{c'}{\sqrt{\|c'\|}} \right) \quad (30)$$

But we can go further and benefit from the specific case that we are dealing with : the data are in a Lie group, $G = SO(n)$, in this section, we will denote the base manifold G to emphasize its group structure, and g is an element of the group. Recall that a Lie group G is a set that is both a smooth manifold and a group, that is this is a manifold in which the group multiplication $G \times G \rightarrow G, (x, y) \mapsto xy$ and the inversion map, $g \rightarrow G, x \mapsto x^{-1}$ are smooth.

Therefore, we can expect a simplification on the previous expressions. The main strength of a Lie group G is that it can be approximated by its Lie algebra \mathfrak{g} , or in other words, the Lie algebra is the best *linear approximation* of a Lie group. The Lie algebra is homeomorphic to the tangent space at the identity. As in [13] we consider only curve that start at the identity: this is because other curve can be reduced to this case by right or left translation. In this settings it is interesting to turn the SRV function into the transported SRV function (TSRV). This is basically the SRV that has been parallel transport to a reference point. Different version has been given, as it can be seen in [9], [13], [54]. They differ by the choice of their reference point. As we study curve in a Lie group, the identity (the starting point of our curve in our case) is a particularly good choice as a reference point. In a Lie group a parallel transport operation can be defined, here again, by the right (or left) translation. This justify that we can take, as suggested in ([13]) a TSRV function of the form:

$$F_{Lie} : \mathcal{C}^\infty([0, 1], G) \longrightarrow SO(n) \times \{q \in \mathcal{C}^\infty([0, 1], \mathfrak{g}), q(t) \neq 0, \forall t \in [0, 1]\} \\ F_{Lie}(c)(t) = (c(0), q(t)) = \left(c(0), \frac{R_{c(t)*}^{-1}(c'(t))}{\sqrt{\|c'(t)\|}} \right) = \left(c(0), \frac{T_c^{c(t) \rightarrow I}(c'(t))}{\sqrt{\|c'(t)\|}} \right) \quad (31)$$

where R is the right translation on the group, $R_{g_1}(g_2) = g_2g_1$, $R_{g*} = T_e R_g$ is the tangent map (the derivative equivalent on manifolds, equivalent notation : dR_g) at the identity, and $\|\cdot\|$ 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 . So a curve is now represented as an element of the tangent bundle $(c(0), q(t)) \in M \times TM$ (recall that q draw a curve in the tangent bundle), and $c(0)$ is the identity element of the Lie group. As mentioned earlier, the TSRV function is the SRV function that has been transported to a certain point, in other words, we use the natural reference point of a Lie group that consists in its identity element to define the parallel transportation, and we apply it to the tangent curve. The inverse of the SRV function is then straightforward: for every $q \in \mathcal{C}^\infty([0, 1], TM)$, there exists a unique curve c such that $F(c) = q$. We have $c(t) = \int_0^t q(r) \parallel q(r) \parallel dr$ where $\parallel \cdot \parallel$ is the norm in $SO(n)$.

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 distance can be found in [30]. The distance on the shape space is used to compare how to curve are intrinsically different. It has been in [37] that the simple L^2 metric on \mathcal{M} given by

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

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. Another type of metric has been investigated, which induced non-vanishing metrics on the shape space: the idea is to add higher order derivatives. These are the *elastic metric*, derived from the Sobolev metric [10], [23], which in the case of an Euclidean space \mathbb{R}^n :

$$g_c^{a,b}(u, v) = \int (a^2 \langle D_l u^N, D_l v^N \rangle + b^2 \langle D_l u^T, D_l v^T \rangle) \parallel c'(t) \parallel dt \quad (32)$$

Where: $\langle \cdot, \cdot \rangle$ is the Euclidean canonical metric on the based manifold ($SO(n)$ in our case), $D_l u = h' / \parallel c' \parallel$, $D_l u^T = \langle D_l u, w \rangle w$, with $w = c' / \parallel c' \parallel$ and $D_l u^N = D_l u - D_l u^T$ this way, $(D_l u^N, D_l u^N)$ defines a mobile frame along the curve c , see figure (3).

Here we are only interesting on the special metric thats has been proposed in [30], and which is an adaptation of the elastic metric for Riemannian manifold:

$$g_c(u, v) = \langle u(0), v(0) \rangle + \int \left(\langle \nabla_l u^N, \nabla_l v^N \rangle + \frac{1}{4} \langle \nabla_l u^T, \nabla_l v^T \rangle \right) \parallel (c'(t)) \parallel dt \quad (33)$$

Where $\langle \cdot, \cdot \rangle$ denotes the Riemannian metric of the original space $SO(n)$, ∇ is the Levi-Civita connection that corresponds to $\langle \cdot, \cdot \rangle$; $\nabla_l u = \frac{1}{\parallel c' \parallel} \nabla_{c'} h$, $\nabla_l u^T = \langle \nabla_l u, w \rangle w$, $w = c' / \parallel c' \parallel$. The computations being done now in a manifold space, the Levi-Civita connection replaced the ordinary derivative of \mathbb{R}^n .

This metric is obtained as the pullback of the canonical L^2 metric in the SRV framework. Recall that if we have a map $\phi : \mathcal{U} \rightarrow \mathcal{V}$ and a one-form (that is a section) α of $T^*\mathcal{V}$, the cotangent bundle of \mathcal{V} , the pullback of α is the one-form *on* \mathcal{U} defined by : $(\phi^* \alpha)_x(X) = \alpha_{\phi(x)}(d\phi(x)(X))$, for all $x \in \mathcal{U}$, $X \in T_x \mathcal{U}$. Here, $\alpha = F$, and $\mathcal{U} = \mathcal{M}$, $\mathcal{V} = TM$ and as we the ambient space is a Lie group we see that the metric is the pullback of the L^2 -metric on the tangent bundle (*i.e.* an inner product on the double tangent bundle TTM) defined by

$$\tilde{g}_c(f, g) = \int_0^1 \langle \nabla_s q^f(0, t), \nabla_s q^g(0, t) \rangle dt \quad (34)$$

where for any $f \in T_c$ a path of curve $s \mapsto c_f(s) \in \mathcal{M}$ is defined, as usually, such that $c_f(0) = c$, $(\partial c_f / \partial s)(0) = f$ (recall that a tangent f vector to a point c can be seen as an equivalence class of infinitesimal deformations of a path that passes through c at time 0). q^f is the TSRV function of c_f . Due to the transportation of the TSRV, the starting point is no longer required in the TSRV domain. Thus, from the definition of a metric on TM , that apply on the space of TSRV functions we can deduce a metric on the space of curves as $g_c(u, v) = \tilde{g}_{F(c)}(T_c F(u), T_c F(v))$, where $T_c R$ denote the tangent map (the derivative) of the TSRV function at c (which carry the information about the starting point of the curve).

What is particular with this metric compared to the classical elastic metric, is that the starting point of the curves intervene explicitly, which is quite important when we deal with periodic curves, as explained before.

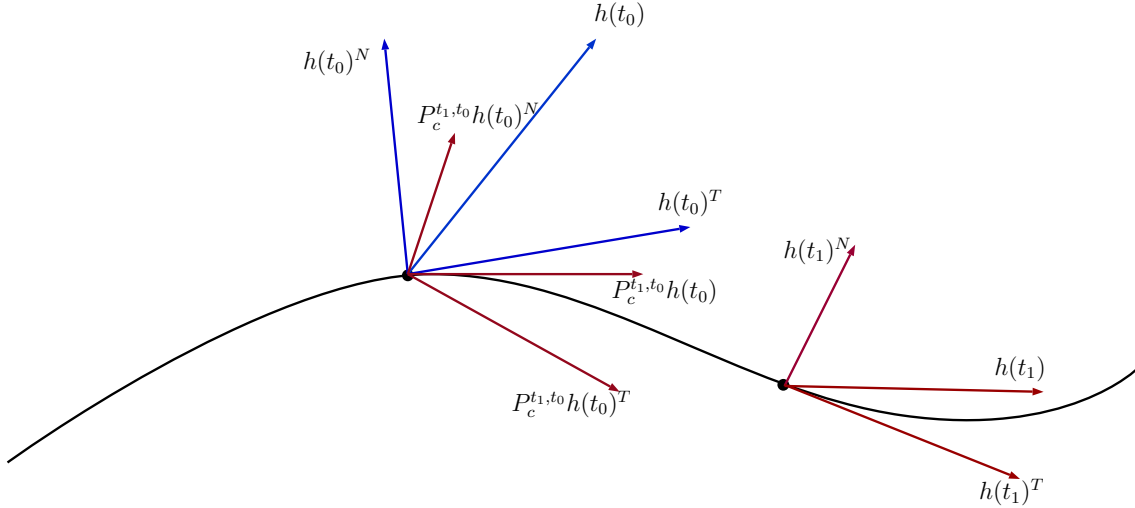


Figure 3: The inner product measures the angle between a frame at a given point and the parallel transport version of this frame at a latter time.

It could be possible to add a term such that $\langle c'_0(0), c'_1(0) \rangle_{TSO(n)}$ for some metric on $SO(n)$ to underline the fact that curves are closed even in their first order derivatives.

One can notice that the SRV formulation of the metric takes a quite simple form. In this framework, the length of a path of curve (notice that it is the length of a path of curve and not the length of a curve in $SO(n)$) is

$$L(c) = \int_0^1 \sqrt{\|x(s)\|^2 + \int_0^1 \|\nabla_{\partial c / \partial s} q(s, t)\|^2 dt} ds \quad (35)$$

Once the geometry has been settled in \mathcal{M} , the geometry of the state-space can be derived from its quotient structure (recall that the space of curve induced a fiber bundle structure above the state-space). Before we have to remember the definition of the decomposition of the tangent bundle as an horizontal and a vertical tangent bundle : given the fiber bundle $\pi : \mathcal{M} \rightarrow \mathcal{S}$ the tangent bundle can be decomposed into a vertical and a horizontal subspace :

$$T\mathcal{M} = \mathcal{H}_{\mathcal{M}} \oplus \mathcal{V}_{\mathcal{M}} \quad (36)$$

with $\mathcal{V}_{\mathcal{M}} = \ker(T_c \pi)$ (T_c is the tangent map), and $\mathcal{H}_{\mathcal{M}} = (\mathcal{V}_{\mathcal{M}})^\perp$ where the orthogonality has to be understood in the sense of the metric defined earlier. This metric is reparametrization invariant, that is constant along the fibers, so it can be proven that the horizontal space $\mathcal{H}_{\mathcal{M}}$ (which concern the fibers) and the tangent space in the shape-space are *isometric* : π is a *Riemannian submersion*. Translated in terms of metrics, we have :

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

where $[g]$ denote the metric *on the shape space*.

Similar result in a different (but still close) context is used in [53], 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 state-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 and c_1 intersect the fiber of c_1 at the a re-parametrized version of c_1 , $c_1 \circ \phi$ which give the distance in the shape space :

$$[d]([c_0], [c_1]) = d_g(c_0, c_1 \circ \phi) \quad (37)$$

where $[d]$ denote the distance in \mathcal{S} , and d_g the distance on the space of curve induced by the aforementioned riemannian metric. Here again a recursive procedure [30] enable the computation of the horizontal geodesic. In the TSRV function formulation, the distance problem is equivalent to an optimization problem :

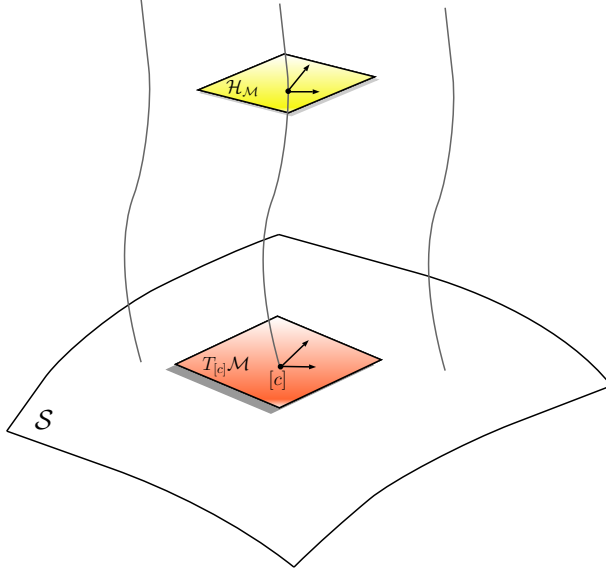


Figure 4: The tangent space $T_{[c]}\mathcal{M}$ at a point $[c]$ in the shape space \mathcal{S} is isomorphic to the horizontal part $\mathcal{H}_{\mathcal{M}}$ of the tangent space at a point on the associated fiber.

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

where $q_i(\cdot) = F(c_i)(\cdot)$. As these curves in the tangent space had been transported according to the same point, which is the identity element, they lived in the same space and can be compared. In [13] the author used a dynamic linear programming to solve this linear optimization problem, they construct a piecewise linear approximation of ϕ . A traditional gradient descent algorithm is also possible.

Last but not least, we have to mention that in a practical situation, the above formula have to be discretized. This is the object of [31]. Formulas 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

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

Concerning the metric on the space of curve, it becomes

$$g_{c_{disc}}(u, v) = \langle u_0, v_0 \rangle + \frac{1}{n} \sum_{i=0}^{n-1} \langle \nabla_{\partial c / \partial s} q^u \left(0, \frac{k}{n} \right), \nabla_{\partial c / \partial s} q^v \left(0, \frac{k}{n} \right) \rangle \quad \forall u, v \in T_{disc}\mathcal{M} \quad (39)$$

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 condition are fulfilled, *i.e.*

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

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 $x_{(k+1)/n}$. Thus the SRV functions that appears in the formula refers 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

$$\nabla_{\partial c / \partial s} q \left(s, \frac{k}{n} \right) = \nabla_{\partial c / \partial s} q_k(s) \quad (40)$$

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 object.

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

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

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 give 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 we are can be exploited for our study to curve on shape spaces. We list some of them in the following

- The geodesics through e (the identity element) are the integral curves $t \mapsto \exp(tu)$, $u \in \mathfrak{g}$, i.e. the one-parameter groups. Also, since 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

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

Of course we have

$$\gamma'(0) = (dL_a)_e(u).$$

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

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 $B : \mathfrak{g} \times \mathfrak{g} \rightarrow \mathbb{C}$ given by $B(u, v) = \text{tr}(ad(u) \circ ad(v))$, where tr denote the trace operator and ad 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 now we assume B to be negative-definite, then $-B$ is an inner product and is adjoint-invariant. So, 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 $Ric(u, v) = -\frac{1}{4}B(u, v)$.

The Lie algebra of $SO(n)$ is the set of skew-symmetric matrices, that is matrices M which verifies $M^T = -M$. The Killing form on $SO(n)$ is given by : $B_{\mathfrak{so}(n)} = (n-2)\text{tr}(XY)$, and due to the skew-symmetry, we have $-B_{\mathfrak{so}(n)} = (n-2)\text{tr}(XY^T)$. Therefore, it induces a bi-invariant metric and the previous formula can be plugged in expression of the metric on the space of curve or on the shape space. To conclude these preliminaries, we see that due to the simpler form of the parallel transportation and of the metric, the distance equations (38) are now easier to handle.

It is now time to give the geodesic equation, relative to our chosen measure. Due to the TSRV, the geodesic equation takes a much simpler form than what can be found in [30] and [31]. The formula can be found in [13]. For the sake of completeness, we give a reformulated proof in Annex 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 (B). So, 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):

$$q = \frac{\partial c / \partial t}{\sqrt{\| \partial c / \partial t \|^2}}$$

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*

$$\nabla_{\partial c / \partial s} (\nabla_{\partial c / \partial s} q(s, t)) (s, t) = 0 \quad \forall s, t \quad (41)$$

Proof. Annex 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:

$$F_{Lie}^{-1}((1-s)F_{Lie}(c_0) + sF_{Lie}(c_1)) \quad (42)$$

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

Finally we can give a slightly modified version of the algorithm of geodesic shooting. As we have seen for example in equation (39) that a piecewise (continuous) geodesic is required to express the metric. This is because the expression are formally constructed with initial conditions (initial point and initial tangent vector). In [13] an exemple is given for curves on $SO(3)$. We give here a simplified version of their construction. Starting from a continuous curve, they discretize it as $\{x_0, \dots, x_n\}$ and then construct a piecewise geodesic c between these points:

$$c(t) = \sum_{k=0}^{n-1} \chi_{k,k+1}(t) \exp((t-k) \log_{c_k}(c_{k+1})) c_k \quad (43)$$

Notice that this geodesic is in the base manifold, and not in the space of curve. Thus, the geodesics are expressed in terms of one-parameter groups. But in our case, we already have a discrete curve. The problem is the number of points that we have. Results exists that show the convergence of the discrete case to the continuous case, but when we deal with positive definite matrice with a certain periodicity, we may have not enough matrices. So, a step has to be added, which is another interpolation step. But in order to have a curve at least \mathcal{C}^1 , and also because distance between the W matrices can be quite high, we interpolate first by spline. There are many way to thought spline interpolation on manifold, 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.

Therefore, our procedure to compare curves on manifold arising from positive definite matrices is the following :

1. **Input** : a set of rotation matrices $\{W_i\}_i$, seen as a partially observation of a periodic trajectory on $SO(n)$.
2. Interpolate with splines between matrices W_i .
3. Truncate the previous curve to obtain a sequence of points $\{x_0, \dots, x_n\}$ with a finer discretization.
4. Interpolate between these points to obtain a piecewise geodesic curve *in the base manifold* $SO(n)$.
5. Apply geodesic shooting to obtain a geodesic path *between two curves*, and eventually between the shapes of the two curves.
6. **Output** : distance between two curve in the manifold defined by the set of curves in $SO(n)$

Formally, we have to add a preliminary step when one deal's with curve that do not starts at the identity. We can either compute the TSRV representation of the curve and then transport this representation to the Lie algebra, or first translate the curve to a curve that starts at the identity. To end this part, we mention that in the general framework of the dilation, we deal with operators of infinite size. These operators are unitary. So, computations are to be adapted: and the inner product on the tangent space is now derived from the Hilbert Schmidt norm :

$$\| A \|_{H.S} = \text{tr}(AA^*) = \sum_{k=1}^{+\infty} \lambda_k(AA^*) \quad (44)$$

where A is a trace class operator, and $\{\lambda_k(AA^*)\}_k$ is the eigenvalues of AA^* . This is a generalization of the Frobenius norm to the infinite dimensional case.

As these operator give rise to the spectral measure of the underlying process, we therefore have a way to compare spectral measures, in a geometrical way that is different from the "classical" use of Fisher information metric or Kullback-Leibler distance in the information geometry context.

4 Conclusion

We have introduced a new vision of stochastic processes through the geometry induced by the dilation. The dilation matrices of a given processes are obtained by a composition of rotations whose angle correspond to the well-known reflexion coefficients. The advantage to work 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 non-stationary, a set of dilation matrices is obtained and it is related to the Kolmogorov decomposition. Rigorously, dilation matrices are infinite dimensional, though we turn them into rotation matrices by truncation. Each of them belongs so 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 focus our attention on the Periodically Correlated (PC) class of non-stationary 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 extent to the space of shapes. It is then possible to compare the shape of curves when the metric is translated to the Lie algebra, achieving therefore a closed-form expression and easy computation.

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 factorization :

$$\begin{pmatrix} X & Y \\ Y^* & Z \end{pmatrix} = \begin{pmatrix} X^{1/2} & 0 \\ Z^{1/2}\Gamma^* & Z^{1/2}D_\Gamma \end{pmatrix} \begin{pmatrix} X^{1/2} & \Gamma Z^{1/2} \\ 0 & D_\Gamma Z^{1/2} \end{pmatrix} \quad (45)$$

Where X and Y are positive matrices. Note that this is a Cholesky factorization-type result. This type of decomposition is used as squared-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 \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 $\Gamma Xh = Yh$, [50]. □

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 chose 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 factorization of positive definite kernel. Note that in the next, $\mathcal{R}(\Gamma)$ will denote the close range of the operator Γ . We first start with a basic case:

Theorem 5 (row contraction). *Let $T = [T_1 \ T_2] \in \mathcal{L}(\mathcal{H}_1 \oplus \mathcal{H}_2, \mathcal{H})$, then $\|T\| \leq 1$ 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*

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

Proof. The proof is a simple application of theorem 4. For the if part, it is obvious that we can take Γ_1 to be T_1 . Then $\|T\| \leq 1$ implies

$$I - TT^* = I - \Gamma_1\Gamma_1^* - T_2T_2^* \geq 0$$

with $D_{\Gamma_1^*}^2 \geq T_2T_2^*$. There exists so 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 factorization, we can write down the defect operator for the whole contraction $T = [T_1 \ T_2]$ [50] to be

$$D_T^2 = \begin{pmatrix} D_{\Gamma_1} & 0 \\ -\Gamma_2^* \Gamma_1 & D_{\Gamma_1} \end{pmatrix} \begin{pmatrix} D_{\Gamma_1} & -\Gamma_1^* \Gamma_2 \\ 0 & D_{\Gamma_1} \end{pmatrix} \quad (46)$$

Therefore, with the theorem 5 we have an operator α such that

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

Similarly,

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

and the general case is

Theorem 6 (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 the form

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

and

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

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 column contraction of the form $T = \begin{pmatrix} T_1 \\ \vdots \\ T_2 \end{pmatrix}$.

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 arise when considering the SRV and not only the TSRV function of a cuve, that is the curve are parametrized by their starting point and their velocity but their starting point 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:*

$$\nabla_{\partial c / \partial s} c(s, 0) + \int_0^1 \mathcal{R}(q(s, t), \nabla_{\partial c / \partial s} q(s, t))(c(s, 0)) dt = 0 \quad \forall s \quad (47)$$

$$\nabla_{\partial c / \partial s} (\nabla_{\partial c / \partial s} q(s, t))(s, t) = 0 \quad \forall s, t \quad (48)$$

Similarly to [30] and [54], 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 (5), to get a clear picture, we have represented a variation of a path of curves with fixed starting and ending points. Though similar, the situation here is a bit different because of the representation of the curve through its SRV function, which we can hardly represent. But the process remains similar. We emphasize the subtle difference with [30]. Here, we work

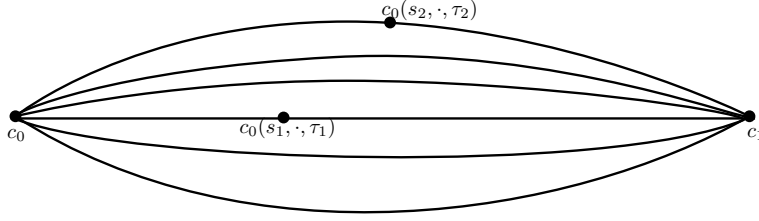


Figure 5: we consider a beam of curves, which consists in a slight modification of the geodesic. The different curves are indexed by τ . The idea is to find which of these curves give the minimal energy to go from c_0 to c_1

directly in the tangent space representation, via the SRV representation, and not with "the whole family" of curves $c(s, t, \tau)$. We denote $\partial_\tau c(s, 0, \tau) = \frac{\partial c(s, 0, \tau)}{\partial \tau}$, 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

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

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

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

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

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

Integrating by parts now, allows to have

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

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

which yields to:

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

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

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

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 due to the fixed starting point which corresponds to the identity element. This simplifies a lot 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. Number 70 in NATO ASI Series, pages 385–395. Springer Berlin Heidelberg, 1991.
- [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(4):595–604, August 2013.
- [3] Gr. Arsene and T. Constantinescu. The structure of the Naimark dilation and Gaussian stationary processes. *Integr equ oper theory*, 8(2):181–204, March 1985.
- [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(6):32–41, nov 2013.
- [5] B. Balaĳi, F. Barbaresco, and A. Decurninge. Information geometry and estimation of Toeplitz covariance matrices. In *2014 Int. Radar Conf.*, pages 1–4. IEEE, oct 2014.
- [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*, pages 124–163. Springer, 2009.
- [7] F. Barbaresco. Geometric Radar Processing based on Fréchet distance: Information geometry versus Optimal Transport Theory. In *2011 12th Int. Radar Symp.*, pages 663–668, 2011.
- [8] F. Barbaresco. Information Geometry of Covariance Matrix: Cartan-Siegel Homogeneous Bounded Domains, Mostow/Berger Fibration and Fréchet Median. In *Matrix Inf. Geom.*, pages 199–255. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013.
- [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*, 34:139–165, 2014.
- [10] M. Bauer, M. Bruveris, and P. W. Michor. Why Use Sobolev Metrics on the Space of Curves. In *Riemannian Computing in Computer Vision*, pages 233–255. Springer, Cham, 2016. DOI: 10.1007/978-3-319-22957-7_11.
- [11] N. H. Bingham. Szego’s Theorem and its Probabilistic Descendants. aug 2011.
- [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*, pages 900–905, 2014.
- [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(6):346, 1996.
- [18] F. Desbouvries. Non-Euclidean geometrical aspects of the Schur and Levinson-Szego algorithms. *IEEE Trans. on Information theory*, 49(8):346, 2003.
- [19] F. Desbouvries. Non-euclidean geometrical aspects of the schur and levinson-szego algorithms. *IEEE Transactions on Information Theory*, 49(8):1992–2003, August 2003.
- [20] S. Dégerine and S. Lambert-Lacroix. Characterization of the partial autocorrelation function of nonstationary time series. *Journal of Multivariate Analysis*, 87(1):46–59, October 2003.

- [21] P. Flandrin. In *Time-Frequency/Time-Scale Analysis*, volume 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*, number 44 in OT 44 Operator Theory: Advances and Applications, pages 497–546. Birkhäuser Basel, 1990. DOI: 10.1007/978-3-0348-7712-1_15.
- [23] A. Gelman, J. B. Carlin, H. S. Stern, D. B. Dunson, A. Vehtari, and D. B. Rubin. *Bayesian data analysis*, volume 2. CRC press Boca Raton, FL, 2014.
- [24] H. L. Hurd and A. Miamee. *Periodically Correlated Random Sequences: Spectral Theory and Practice*. Wiley Series in Probability and Statistics. Wiley, 2007.
- [25] X. Jiang, L. Ning, and T.T. Georgiou. Distances and Riemannian Metrics for Multivariate Spectral Densities. *IEEE Trans. Automat. Contr.*, 57(7):1723–1735, jul 2012.
- [26] T. Kailath and A. M. Bruckstein. Naimark Dilations, State-Space Generators and Transmission Lines. In R. G. Douglas, C. M. Pearcy, B. Sz.-Nagy, F.-H. Vasilescu, D. Voiculescu, and Gr. Arsene, editors, *Advances in Invariant Subspaces and Other Results of Operator Theory*, number 17 in Operator Theory: Advances and Applications, pages 173–186. Birkhäuser Basel, 1986. DOI: 10.1007/978-3-0348-7698-8_13.
- [27] E.K. Kalunga, S. Chevallier, Q. Barthélemy, K. Djouani, Y. Hamam, and E. Monacelli. From Euclidean to Riemannian Means: Information Geometry for SSVEP Classification. pages 595–604. Springer, Cham, oct 2015.
- [28] S. Kolouri, S.R. Park, M. Thorpe, D. Slepcev, and G.K. Rohde. Optimal Mass Transport: Signal processing and machine-learning applications. *IEEE Signal Process. Mag.*, 34(4):43–59, jul 2017.
- [29] S. Lambert-Lacroix. Extension of Autocovariance Coefficients Sequence for Periodically Correlated Processes. *Journal of Time Series Analysis*, 26(3):423–435, May 2005.
- [30] A. Le Brigant. Computing distances and geodesics between manifold-valued curves in the SRV framework. *arXiv preprint arXiv:1601.02358*, 2016.
- [31] A. Le Brigant. A discrete framework to find the optimal matching between manifold-valued curves. *arXiv preprint arXiv:1703.05107*, 2017.
- [32] A Makagon and H Salehi. Notes on infinite dimensional stationary sequences. In S Cambanis and A Weron, editors, *Probability Theory on Vector Spaces IV*, volume 1391 of *Lecture Notes in Mathematics*, pages 200–238. Springer Berlin Heidelberg, 1989.
- [33] J.H. Manton. A Primer on Stochastic Differential Geometry for Signal Processing. *IEEE J. Sel. Top. Signal Process.*, 7(4):681–699, aug 2013.
- [34] P. Masani. Dilations as Propagators of Hilbertian Varieties. *SIAM J. Math. Anal.*, 9(3):414–456, June 1978.
- [35] A.G. Miamee. Spectral dilation of $L(B,H)$ -valued measures and its application to stationary dilation for Banach space valued processes. *Indiana Univ. Math. Journal*, 38:841–860, 1989.
- [36] A.G. Miamee. Periodically Correlated Processes and Their Stationary Dilations. *SIAM Journal on Applied Mathematics*, 50(4):1194–1199, 1990.
- [37] P. W. Michor and D. Mumford. Vanishing geodesic distance on spaces of submanifolds and diffeomorphisms. *Doc. Math.*, 10:217–245, 2005.
- [38] N. Lipeng, J. Xianhua, and T.T Georgiou. Geometric methods for structured covariance estimation. In *2012 Am. Control Conf.*, pages 1877–1882. IEEE, jun 2012.
- [39] A. Napolitano. Cyclostationarity: New trends and applications. *Signal Processing*, 120:385–408, 2016.
- [40] X. Pennec. Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements. *Journal of Mathematical Imaging and Vision*, 25(1):127–154, 2006.

- [41] X. Pennec, P. Fillard, and N. Ayache. A Riemannian framework for tensor computing. *International Journal of Computer Vision*, 66(1):41–66, 2006.
- [42] A. H. Sayed, T. Constantinescu, and T. Kailath. Recursive Construction of Multichannel Transmission Lines with a Maximum Entropy Property. In *Codes, Graphs, and Systems*, pages 259–290. Springer, 2002.
- [43] B. Simon. CMV matrices: Five years after. *Journal of Computational and Applied Mathematics*, 208(1):120–154, November 2007.
- [44] B. Simon. *Orthogonal Polynomials on the Unit Circle Part 1 and Part 2*, volume 54. American Mathematical Society, 2009.
- [45] S T Smith. Covariance, subspace, and intrinsic Cramér-Rao bounds. *IEEE Trans. on Signal Process.*, 53(5):1610–1630, 2005.
- [46] 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(7):1415–1428, July 2011.
- [47] K. Sun and S. Marchand-Maillet. An information geometry of statistical manifold learning, 2014.
- [48] B. Sz.-Nagy, C. Foias, H. Bercovici, and L. Kérchy. *Harmonic Analysis of Operators on Hilbert Space*. Springer New York, New York, NY, 2010.
- [49] D. Timotin. Prediction theory and choice sequences: an alternate approach. In *Advances in invariant subspaces and other results of operator theory*, pages 341–352. Springer, 1986.
- [50] M. C. Tseng. Contractions, Matrix Parametrizations, and Quantum Information. *arXiv:quant-ph/0610259*, October 2006. arXiv: quant-ph/0610259.
- [51] M. C. Tseng and V. Ramakrishna. Dilation Theoretic Parametrizations of Positive Matrices with Applications to Quantum Information. *arXiv:quant-ph/0610021*, October 2006. arXiv: quant-ph/0610021.
- [52] L. Yang, M. Arnaudon, and F. Barbaresco. Riemannian median, geometry of covariance matrices and radar target detection. In *Radar Conference (EuRAD), 2010 European*, pages 415–418, September 2010.
- [53] Yili L. and K. M. Wong. Riemannian Distances for Signal Classification by Power Spectral Density. *IEEE Journal of Selected Topics in Signal Processing*, 7(4):655–669, August 2013.
- [54] 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.