Hamburger Beiträge
zur Angewandten Mathematik

Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction

Mijail Guillemard and Armin Iske

Nr. 2009-16
November 2009

Curvature Analysis of Frequency Modulated Manifolds in Dimensionality Reduction

Mijail Guillemard* Armin Iske†

Department of Mathematics, University of Hamburg, Germany

Abstract

Recent advances in the analysis of high-dimensional signal data have triggered an increasing interest in geometry-based methods for nonlinear dimensionality reduction (NDR). In many applications, high-dimensional datasets typically contain redundant information, and NDR methods are important for an efficient analysis of their properties. During the last few years, concepts from differential geometry were used to create a whole new range of NDR methods. In the construction of such geometry-based strategies, a natural question is to understand their interaction with classical and modern signal processing tools (convolution transforms, Fourier analysis, wavelet functions). In particular, an important task is the analysis of the incurred geometrical deformation when applying signal transforms to the elements of a dataset. In this paper, we propose the concepts of frequency modulation maps and modulation manifolds for the construction of particular datasets relevant in signal processing and NDR. Moreover, we design a numerical algorithm for analyzing geometrical properties of the modulation manifolds, with a particular focus on their scalar curvature. Finally, in our numerical examples, we apply the resulting geometry-based analysis algorithm to two model problems, where we present geometrical and topological effects of relevance in manifold learning.

Keywords: Nonlinear dimensionality reduction, manifold learning, signal processing, Fourier and wavelet analysis, numerical differential geometry.

1 Introduction

During the last decade, novel concepts for *nonlinear dimensionality reduction* (NDR) have gained enormous popularity due to the ever increasing complexity of new challenging application problems. In the design of these modern tools, special emphasis is placed on geometrical aspects, where concepts from differential geometry play an important role [2,7,8,13]. The geometry-based approach of NDR can be viewed as a complementary strategy to statistical oriented methods from machine learning and data mining [5].

*guillemard@math.uni-hamburg.de

†iske@math.uni-hamburg.de

To briefly describe the basic problem of NDR and manifold learning, suppose we are given a dataset $X = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$ lying in a high-dimensional Euclidean space, where X is assumed to be sampled from a submanifold $\mathcal{M}$ of $\mathbb{R}^n$, i.e., $X \subset \mathcal{M} \subset \mathbb{R}^n$. Moreover, we assume that the dimension of $\mathcal{M}$ is much smaller than the dimension of the ambient space, i.e., $\dim(\mathcal{M}) \ll n$. The primary objective of manifold learning is to construct a low-dimensional representation of X which can be used to efficiently visualize and analyze its geometrical properties.

For many examples of datasets $X = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$, each element $x_i \in X$ can be considered as a signal that may be analyzed through a transformation map T, defined via convolution transforms, Fourier analysis, or wavelet functions. Therefore, from a manifold learning perspective, it is quite natural to analyze the geometrical deformation between X and $T(X) = \{T(x_i)\}_{i=1}^m$, as being incurred by T, or (if the transformation T is used in a preprocessing step prior to the application of a dimensionality projection map P) being incurred by a composition $P \circ T$ of a transformation T and a projection P.

To investigate these problems, we analyze in this paper a particular class of datasets X and manifolds $\mathcal{M}$ which are generated by frequency modulation maps. Moreover, we propose a numerical method, which serves to analyze various geometrical properties of the input datasets. We remark that the idea of using modulation manifolds is inspired by concepts of frequency modulation for signal transmission in engineering domains. To gain further insight into their geometrical properties, and to reduce the high complexity of necessary algebraic operations involved, we work with numerical approximations to construct basic geometric data such as metric and curvature tensors. By using modulation manifolds we can design examples of low-dimensional data sets embedded in high dimensional spaces, which are relevant in applications of signal and image processing. One peculiarity of these examples is that they lead (in the same spirit as the Swiss role data) to critical test case scenarios, where classical linear projections, such as principal component analysis (PCA) and multidimensional scaling (MDS), are outperformed by more recent nonlinear methods, such as isomap, local tangent space alignment (LTSA), Riemannian normal coordinates (RNC), to mention but a few.

The outline of this paper is as follows. In the following Section 2, basic features of manifold learning and dimensionality reduction are reviewed, where also a short discussion concerning the interaction of dimensionality reduction maps and signal transforms is provided. Then, in Section 3 the concept of modulation maps between manifolds is explained in detail. This is followed by a discussion on metric and curvature tensors, as they are used for measuring geometric deformations of modulation manifolds. In Section 4, numerical examples are provided, where the geometric distortion of selected manifolds (sphere and torus surfaces) is illustrated in situations where frequency modulation maps and dimensionality reduction methods are combined. In the numerical examples, we also compare standard PCA techniques with modern (nonlinear) dimensionality reduction strategies. Various interesting effects concerning the geometrical and topological distortion of the underlying manifolds are presented in Section 4.

2 Manifold Learning in Dimensionality Reduction

Let $\mathcal{M} \subset \mathbb{R}^n$ denote a smooth compact Riemannian submanifold of a high-dimensional space $\mathbb{R}^n$, where the dimension of $\mathcal{M}$ is much smaller than that of the ambient space, i.e., $\mathcal{M} \subset \mathbb{R}^n$ with $p = \dim(\mathcal{M}) \ll n$. The primary goal of manifold learning is to construct a low-dimensional representation Ω of $\mathcal{M}$, with assuming the existence of a diffeomorphism $\mathcal{A} : \Omega \subset \mathbb{R}^d \to \mathcal{M} \subset \mathbb{R}^n$, $d \ll n$. One relevant task that we wish to address here is the analysis of a discrete sample of $\mathcal{M}$, given by a finite scattered dataset $X = \{x_i\}_{i=1}^m \subset \mathcal{M}$. In this problem, one aims at the construction of a low-dimensional representation $Y = \{y_i\}_{i=1}^m \subset \Omega \subset \mathbb{R}^d$ sharing similar geometrical properties with X. This in turn requires a suitable smooth map $E : \mathcal{M} \to \Omega$, whose only input is the dataset $X \subset \mathcal{M}$. Due to the Whitney embedding theorem (which states that any connected smooth p-dimensional manifold can smoothly be embedded in $\mathbb{R}^{2p+1}$), one basic condition in this problem is $2p + 1 \leq d \leq n$, see [1]. Throughout this paper, we use the term *manifold* to denote a compact smooth connected manifold embedded in the Euclidean space $\mathbb{R}^n$.

Now a crucial requirement in manifold learning is to ensure conditions under which the finite sampling $X = \{x_i\}_{i=1}^m$ is dense enough for recovering the geometrical properties of $\mathcal{M}$. These conditions have been investigated over the last years [9]. A main concept is the *condition number* $1/\tau$ of the manifold which encodes local and global curvature properties of $\mathcal{M}$. The condition number can be related to the *medial axis* of $\mathcal{M}$, which is defined as the closure of the set

$$G = \{x \in \mathbb{R}^n \ : \exists\, p, q \in \mathcal{M}, p \neq q, \text{ with } d(x, \mathcal{M}) = \|x - p\| = \|x - q\|\}.$$

By using the medial axis of the manifold, we have

$$\tau = \inf_{p \in \mathcal{M}} d(p, \overline{G}).$$

With these concepts, the following result, relating the sampling of the manifold with its homological reconstruction, is discussed in [9].

Proposition 2.1 *Let $\mathcal{M}$ be a compact Riemannian submanifold of $\mathbb{R}^n$ and $X = \{x_i\}_{i=1}^m \subset \mathbb{R}^n$ a finite $\epsilon/2$-dense collection in $\mathcal{M}$, i.e., for each $p \in \mathcal{M}$, there is an $x \in X$ satisfying $\|p - x\|_{\mathbb{R}^n} < \epsilon/2$. Then for any $\epsilon < \sqrt{3/5}\tau$, we have that $U = \bigcup_{x \in X} B_\epsilon(x)$ deformation retracts to $\mathcal{M}$, and therefore the homology of U equals the homology of $\mathcal{M}$.* ■

A series of additional important developments concerning conditions for efficient sampling of manifolds can be found in [10, 11].

2.1 Application Examples

Relevant applications for our investigations are in time-frequency analysis, where a signal is segmented in time consecutive sections, which are then Fourier transformed. The typical context can be described, for instance, in the short term Fourier transform (STFT) for functions $f \in L^2([0, 1])$ with a window function g,

$$\mathcal{G}_g f(b, \omega) = < f, g_{b,\omega} > = \int_0^1 f(t)\overline{g_{b,\omega}(t)}\, dt \qquad \text{where} \quad g_{b,\omega}(t) = g(t - b)e^{2\pi i \omega t}.$$

Now, in the generation of the dataset X, we consider using a bandlimited signal $f \in L^2([0,1])$ and a segmentation of its domain in such a way that small consecutive signal patches are analyzed, as performed in STFT or wavelet analysis. For instance, the set of signal patches can be defined as a dataset of the form

$$X_f = \{x_i\}_{i=1}^m \qquad \text{for} \quad x_i = (f(t_{k(i-1)+j}))_{j=0}^{n-1} \in \mathbb{R}^n$$

for $k \in \mathbb{N}$ being a fixed *hop-size*. Here, the regular sampling grid $\{t_\ell\}_{\ell=0}^{km-k+n-1} \subset [0,1]$ is constructed with considering the Nyquist-Shannon theorem for f. Notice that X_f may be embedded in a very high-dimensional ambient space $\mathbb{R}^n$, although the dimension of X_f itself may be small. For instance, in audio analysis, for 44kHz signals, $n = 1024$ is commonly used, and therefore customized dimensionality reduction methods could be of vital interest. With this particular scheme, the STFT of f can be interpreted as a transformation of the set X by taking the (windowed) Fourier transform of each x_i.

In this paper we regard X as a geometrical object in the context of manifold learning, a strategy that might provide new information when studying the function f. We remark that the approach taken here essentially differs from traditional nonlinear time-series analysis, in the sense that we don't consider any time-delays and embedding dimensions as used in phase space representations [4].

A second family of examples (similar to the spirit of time-frequency analysis) arises in image processing. One strategy would be to consider a dataset $X = \{x_i\}_{i=1}^m$ constructed from an grayscale image $F : [0,1]^2 \to [0,1]$, along with a finite covering of small squares $O_i \subset [0,1]^2$, centered at pixels positions $\{k_i\}_{i=1}^m \subset [0,1]^2$. As in the previous situation, when considering band-limited images, the domain $[0,1]^2$ can be sampled uniformly and the dataset can then be defined as

$$X = \{F(O_i) \in \mathbb{R}^n\}_{i=1}^m,$$

where n is the size of the squares O_i, and m denotes the number of pixels k_i. As before, our aim is to analyze the geometry of the image data X to gain useful information about the properties of the image F.

As already explained in the outset of the introduction, it is desirable to work with analysis methods that combine signal processing transforms with dimensionality reduction methods. In this case, the basic objects are the manifold $\mathcal{M}$, the data samples $X = \{x_i\}_{i=1}^m$ taken from $\mathcal{M}$, and a diffeomorphism $\mathcal{A} : \Omega \to \mathcal{M}$, where Ω is the low-dimensional copy of $\mathcal{M}$ to be reconstructed via dimensionality reduction. In this case, the only algorithmic input is the dataset X, but with the assumption that we can reconstruct topological information from $\mathcal{M}$ with X in the spirit of Proposition 2.1. Another basic object in our scheme is a signal processing map $T : \mathcal{M} \to \mathcal{M}_T$, which may be based on Fourier analysis, wavelet transforms, or convolution filters, together with the resulting set $\mathcal{M}_T = \{T(p), p \in \mathcal{M}\}$ of transformed data. The following diagram shows the basic situation of our analysis.

$$\begin{array}{ccccc} \Omega \subset \mathbb{R}^d & \xrightarrow{\mathcal{A}} & \mathcal{M} \subset \mathbb{R}^n & & \\ & & \downarrow T & & \\ \Omega' \subset \mathbb{R}^d & \xleftarrow[P]{} & \mathcal{M}_T \subset \mathbb{R}^n & & \end{array}$$

The final component constructs an approximation to Ω, denoted as $\Omega' = P(\mathcal{M}_T)$, by using a dimensionality reduction map P. The characteristics of Ω and Ω' may differ depending on the dimensionality reduction technique, but the main objective is to construct Ω', so that basic geometrical and topological properties of Ω are recovered. In this paper, we will use modulation maps for the embedding $\mathcal{A}$ to study geometrical and topological effects being incurred by particular dimensionality reduction projections $P : \mathcal{M} \to \Omega'$.

3 Modulation Maps and Curvature Distortion

This section is devoted to a particular construction of manifolds $\mathcal{M}$ and diffeomorphisms $\mathcal{A}$ based on modulation maps. Modulation techniques are well-known engineering and telecommunication procedures used to transmit information by varying the frequency of a carrier signal. A main property of these techniques is the simultaneous transmission of different information by using different frequency bands that can conveniently be separated with special convolution filters. Motivated by this, we want to analyze, from a geometrical point of view, a frequency modulation map $\mathcal{A} : \Omega \to \mathcal{M}$, where $\mathcal{M}$ represents the carrier signals modulated by Ω, which is the information content to be transmitted. Rather than analyzing an actual engineering modulation technique, we use this modulation map for the construction of manifolds that are relevant in dimensionality reduction and signal analysis. The domain Ω is, in our particular situation, a manifold, and so its structural content, transmitted via $\mathcal{A}$, needs to be extracted from $\mathcal{M}$.

3.1 Frequency Modulation Maps between Manifolds

A *modulation map* between two manifolds, $\Omega \subset \mathbb{R}^d$ and $\mathcal{M} \subset \mathbb{R}^n$, of equal dimension is a mapping $\mathcal{A} : \Omega \subset \mathbb{R}^d \to \mathcal{M} \subset \mathbb{R}^n$ of the form

$$\mathcal{A}_\alpha(t_i) = \sum_{k=1}^{d} \phi_k(\alpha_k t_i + b_k), \quad \alpha = (\alpha_1, \ldots, \alpha_d) \in \Omega, \quad \{t_i\}_{i=1}^n \subset [0,1],$$

for some fixed functions ϕ_k. We use a uniformly spaced finite sampling $\{t_i\}_{i=1}^n \subset [0,1]$ (as justified with the Nyquist-Shannon sampling theorem), since we work with bandlimited functions $\mathcal{A}_\alpha$.

An important example of modulation maps are frequency modulation functions defined with a standard trigonometrical basis. For $\phi(t) = \sin(t)$, the corresponding modulation map $\mathcal{A}$ between Ω and $\mathcal{M}$ is referred to as a *frequency based modulation map*. We will consider a particular case of frequency modulation maps by using the concept of band separation in the following sense. A frequency modulation map $\mathcal{A}$ has *separated frequency bands* if it can be written, for a given $\alpha^0 \in \mathbb{R}^d$, as

$$\mathcal{A}_\alpha(t_i) = \sum_{k=1}^{d} \sin((\alpha_k^0 + \gamma\alpha_k)t_i + b_k),$$

where the bands B_k are *separated*, i.e., $B_k \cap B_j = \emptyset$, for all $k \neq j$, with

$$B_k = \{\alpha_k^0 + \gamma\alpha_k \in \mathbb{R},\ \alpha = (\alpha_1, \ldots, \alpha_{k-1},\ \alpha_k, \alpha_{k+1}, \ldots, \alpha_d) \in \Omega\}.$$

In other words, we apply an affine transform to the domain $\Omega \subset \mathbb{R}^d$, with shift α^0 and scale γ, so that the coordinates of the vectors in the resulting set $\alpha^0 + \gamma\Omega$ will not share common values. In this case the frequency content introduced by each coordinate $\alpha \in \Omega$ will not overlap in $\mathcal{A}_\alpha$.

3.2 Metric Tensor and Curvature Distortion

Now we wish to analyze the geometrical deformation between Ω, $\mathcal{M}$ and Ω', as incurred by the modulation map $\mathcal{A} : \Omega \to \mathcal{M}$ and the dimensionality reduction transformation $P : \mathcal{M} \to \Omega'$. We measure the geometrical deformation by the manifolds' scalar curvature distortion. In this section, we explain the basic ingredients for computing the scalar curvature S of a Riemannian manifold $\mathcal{M}$. As a starting point we regard a metric tensor field for $\mathcal{M}$ [6, 12], being defined for a particular system of local coordinates $(\theta_i, \ldots, \theta_k)$, as

$$g_{ij}(x) = g_{ij}(\theta_i, \ldots, \theta_k) =< \partial_i, \partial_j > .$$

An invariant of a Riemannian manifold with respect to isometries are its sectional curvatures. This concept generalizes the notion of Gaussian curvature to setting of 2-manifolds, and is defined as

$$K_{\mathcal{M}} = \frac{< R(X,Y)Y, X >}{\|X\|^2\|Y\|^2 - < X, Y >^2},$$

for the *curvature tensor* R, defined for a triple of smooth vector fields X, Y, Z as

$$R(X,Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z.$$

We recall that the affine connection (a *Levi-Cevita connection* in our situation) is a bilinear map

$$\nabla : C^\infty(\mathcal{M}, T\mathcal{M}) \times C^\infty(\mathcal{M}, T\mathcal{M}) \to C^\infty(\mathcal{M}, T\mathcal{M})$$

that can be expressed with the Christoffel symbols defined by the relation

$$\nabla_{\partial_i} \partial_j = \sum_{k=1}^{n} \Gamma_{ij}^k \partial_k.$$

The Christoffel symbols can be described with respect to the metric tensor via

$$\Gamma_{ij}^k = \frac{1}{2} \sum_{\ell=1}^{m} \left(\frac{\partial g_{j\ell}}{\partial x_i} + \frac{\partial g_{i\ell}}{\partial x_j} - \frac{\partial g_{ij}}{\partial x_\ell} \right) g^{\ell k}. \tag{1}$$

An explicit formula for the curvature tensor is given in terms of the Christoffel symbols as follows (the 1,3 curvature tensor)

$$R^\ell{}_{ijk} = \sum_{h=1}^{m} (\Gamma_{jk}^h \Gamma_{ih}^\ell - \Gamma_{ik}^h) + \frac{\partial \Gamma_{jk}^\ell}{\partial \theta_i} - \frac{\partial \Gamma_{ik}^\ell}{\partial \theta_j}. \tag{2}$$

We use the tensor contractions

$$R_{ijk\ell} = \sum_{h=1}^{m} R_{ijk}^h g_{\ell h}, \qquad R_{ij} = \sum_{k,\ell=1}^{m} g^{k\ell} R_{kij\ell} = \sum_{k=1}^{m} R_{kij}^k \tag{3}$$

for intermediate computations. The scalar curvature is computed together with Gaussian curvature, which for the case of two dimensional manifolds differs by a factor 2, as

$$S = \sum_{i,j=1}^{m} g^{ij} R_{ij}. \tag{4}$$

Now, in Riemannian geometry two diffeomorphic manifolds, Ω and $\mathcal{M}$ (related with a diffeomorphism $\mathcal{A} : \Omega \to \mathcal{M}$), share the same geometrical properties when considering the *pull-back* of a metric tensor in $\mathcal{M}$, or the *push-forward* of a metric tensor in Ω. But in our particular situation, we are considering manifolds embedded in linear spaces (i.e., $\Omega \subset \mathbb{R}^d$ and $\mathcal{M} \subset \mathbb{R}^n$). Our strategy for introducing the concept of curvature distortion is to compare the geometries of Ω and $\mathcal{M}$ generated by their corresponding first fundamental forms, which are particular metrics induced by their ambient space.

3.3 Metric Tensor for Modulation Manifolds

We now compute a metric tensor in $\mathcal{M}$ in order to analyze the geometrical deformation on manifold Ω as incurred by the application of a modulation map $\mathcal{A} : \Omega \subset \mathbb{R}^d \to \mathcal{M} \subset \mathbb{R}^n$. The resulting metric tensor can then be used for computing the curvature tensor and the corresponding scalar curvature, which will be used as a measure for the geometric deformation. Our strategy is to consider a parametrization of Ω and to compute the metric tensor generated from the ambient space $\mathbb{R}^n$. In particular, we use the first fundamental form with respect to the given parametrization. The resulting formula (5) follows by direct computation, as explained in the following proposition, where we compute the first fundamental form of a modulated manifold $\mathcal{M}$.

Proposition 3.1 *Let $\mathcal{M}$ be a manifold constructed from a diffeomorphic modulation map $\mathcal{A} : \Omega \subset \mathbb{R}^d \to \mathcal{M} \subset \mathbb{R}^n$,*

$$\mathcal{A}_\alpha(t_i) = \sum_{k=1}^{d} \phi_k(\alpha_k t_i + b_k) \qquad \textit{where } \alpha = (\alpha_1, \ldots, \alpha_d) \in \Omega \textit{ and } \{t_i\}_{i=1}^n \subset [0,1],$$

and $\{\alpha_j(\theta_1, \ldots, \theta_p)\}_{j=1}^d$ be a parametrization of Ω with $p = \dim(\mathcal{M}) = \dim(\Omega)$. The first fundamental form of $\mathcal{M}$ constructed from this parametrization is given by

$$g_{sr} = \sum_{\ell=1}^{n} t_\ell^2 \sum_{p,q=1}^{d} \left(\frac{d\phi_p}{dt}(\alpha_p t_\ell + b_p) \frac{d\phi_q}{dt}(\alpha_q t_\ell + b_q) \frac{\partial \alpha_p}{\partial \theta_s} \frac{\partial \alpha_q}{\partial \theta_r} \right). \tag{5}$$

Proof: This follows as a direct computation of the Jacobian of the composition $\mathcal{A} \circ \alpha$. The Jacobian with respect to parametrization $\alpha_j(\theta_1, \ldots, \theta_d)$ of Ω is given by

$$J_{\mathcal{A}} = \left(\frac{\partial \mathcal{A}_\ell}{\partial \theta_i} \right)_{\ell,i} \qquad \begin{aligned} \frac{\partial \mathcal{A}_\ell}{\partial \theta_i} &= \frac{\partial}{\partial \theta_i} \left(\sum_{j=1}^{d} \phi_j(\alpha_j t_\ell + b_j) \right) \\ &= \sum_{j=1}^{d} \frac{d\phi_j}{dt}(\alpha_j t_\ell + b_j) t_\ell \frac{\partial \alpha_j}{\partial \theta_i} \end{aligned}$$

The first fundamental form (metric tensor) of $\mathcal{M}$ is given by

$$\begin{aligned}\left(J_{\mathcal{A}}^T J_{\mathcal{A}}\right)_{s,r} &= \sum_{\ell=1}^{n}\left(\sum_{j=1}^{d}\frac{\partial\phi_j}{\partial t}(\alpha_j t_\ell + b_j)t_\ell\frac{\partial\alpha_j}{\partial\theta_s}\right)\left(\sum_{j=1}^{d}\frac{\partial\phi_j}{\partial t}(\alpha_j t_\ell + b_j)t_\ell\frac{\partial\alpha_j}{\partial\theta_r}\right)\\ &= \sum_{\ell=1}^{n}\sum_{p,q=1}^{d}\left(\frac{d\phi_p}{dt}(\alpha_p t_\ell + b_p)t_\ell\frac{\partial\alpha_p}{\partial\theta_s}\frac{d\phi_q}{dt}(\alpha_q t_\ell + b_q)t_\ell\frac{\partial\alpha_q}{\partial\theta_r}\right)\\ &= \sum_{\ell=1}^{n}t_\ell^2\sum_{p,q=1}^{d}\left(\frac{d\phi_p}{dt}(\alpha_p t_\ell + b_p)\frac{d\phi_q}{dt}(\alpha_q t_\ell + b_q)\frac{\partial\alpha_p}{\partial\theta_s}\frac{\partial\alpha_q}{\partial\theta_r}\right)\end{aligned}$$

As g_{sr} is given by $\left(J_{\mathcal{A}}^T J_{\mathcal{A}}\right)_{s,r}$, we obtain the resulting equation (5). ■

The expression in equation (5) will play a crucial role in our following discussion, but due to its complexity (even for quite simple examples as a sphere or a torus), our strategy is to use a numerical framework for illustrating its properties by using the related curvature tensors. Another reason for taking a numerical approach is to build a flexible scheme that can handle arbitrary two dimensional frequency modulated manifolds $\mathcal{M}$ defined by a finite scattered dataset $X = \{x_i\}_{i=1}^m \subset \mathcal{M}$.

3.4 Numerical Computation of Curvature Tensors

The next step is to combine equation (5) with the computation of the curvature tensors in equations (3) and (4) for describing how Ω is geometrically deformed under the mapping $\mathcal{A} : \Omega \to \mathcal{M}$. In the following computations, we focus on the particular case of two dimensional manifolds embedded in a three dimensional space, namely, $d = 3, \dim(\Omega) = 2$, with $\Omega \subset \mathbb{R}^3$. In order to compute the metric tensor as described in equation (5), the main inputs are the functions $\{\phi_j\}_{j=1}^3$ and the parametrization $\{\alpha_j(\theta_1, \theta_2)\}_{j=1}^3$ of Ω, which is then used to construct the Jacobian components $\frac{\partial\alpha_p}{\partial\theta_s}$ and $\frac{\partial\alpha_q}{\partial\theta_r}$. The following algorithm describes the basic steps for computing the scalar curvature S and the metric tensors g_{ij} of the modulated manifold $\mathcal{M}$.

Algorithm 3.1 (Curvature and Metric Tensors of Manifolds.)

Input:
(a) *Parametrization* $\alpha = (\alpha_j(\theta_1, \theta_2))_{j=1}^3$ *of* Ω*;*
(b) *Functions* $\{\phi_j\}_{j=1}^3$ *generating the map* $\mathcal{A}$*.*

(1) *Compute the Jacobian matrices* J_α*;*

(2) *Compute the metric tensor* g_{ij} *via equation (5);*

(3) *Compute the Christoffel symbols* Γ_{ij}^k *via equation (1);*

(4) *Compute the tensors* $R_{ijk\ell}$*,* R_{ij} *via equations (3);*

(5) *Compute the scalar curvature* S *via equation (4).*

Output: *Scalar curvature* S *of* $\mathcal{M}$*.*

4 Numerical Experiments

We have implemented the proposed method, Algorithm 3.1, for computing the metric and curvature tensors of frequency modulated manifolds. The corresponding Matlab code is available through `www.math.uni-hamburg.de/home/guillemard/curvature/`. In this section, numerical examples concerning two different test cases are presented. In the two test cases, we decided to use the sphere, $\Omega = \mathbb{S}^2$, and the torus, $\Omega = \mathbb{T}^2$, to illustrate how the scalar curvature is modified under modulation maps and dimensionality reduction projections.

4.1 Frequency Modulation for a Sphere

In this first numerical example, we let the manifold Ω be given by the unit sphere, i.e., $\Omega = \mathbb{S}^2 \subset \mathbb{R}^3$, and the modulation map $\mathcal{A} : \mathbb{S}^2 \subset \mathbb{R}^3 \to \mathcal{M} \subset \mathbb{R}^{256}$ is defined as

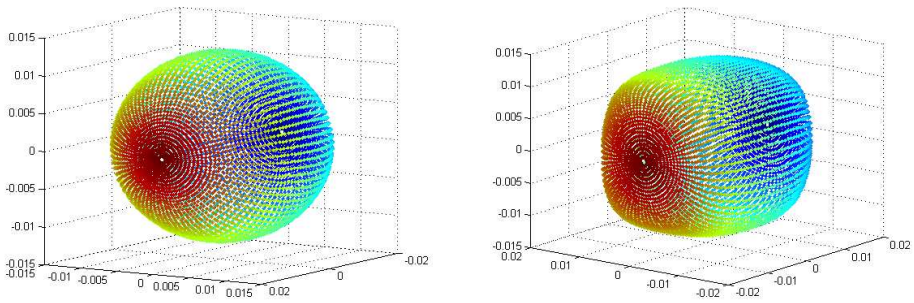
$$\mathcal{A}_\alpha(t_i) = \sum_{k=1}^{3} \sin((\alpha_k^0 + \gamma\alpha_k)t_i + b_k),$$

where

$$\begin{aligned} \alpha_1(u,v) &= \cos(v)\cos(u), \\ \alpha_2(u,v) &= \cos(v)\sin(u), \\ \alpha_3(u,v) &= \sin(v). \end{aligned}$$

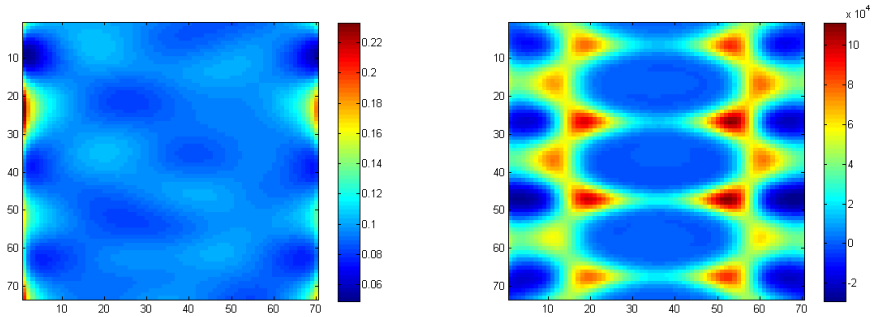
Here, we use a finite and regular distribution of values $u \in [0, 2\pi]$, $v \in [0, \pi]$, and $\{t_i\}_{i=1}^{256} \subset [0, 1]$. We apply shifting and scaling to the manifold $\mathbb{S}^2$, so that the frequency positions are given by the coordinates $(\alpha_0^1, \alpha_0^2, \alpha_0^3)$ and the scaling factor γ. We use these parameters to obtain a separation of the frequency bands as described in Subsection 3.1. The scaling factor γ gives the spreading of each frequency band (bandwidth). A main observation of the following experiments is that the geometrical deformation depends primarily on the parameter γ. A graphical display of the manifolds $\mathbb{S}^2$ and $\mathcal{M}$ is presented in Figure 1, where the sphere $\mathbb{S}^2$ is compared with a three dimensional PCA projection of $\mathcal{M}$ (denoted as $P(\mathcal{M})$). The PCA projection $P(\mathcal{M})$ of $\mathcal{M} \subset \mathbb{R}^n$, $n = 256$, produces a significant geometrical deformation. One objective of the following analysis is to measure this distortion. It can be observed experimentally that an increase in the scale factor γ corresponds to a more pronounced cubic shape deformation, as shown in Figure 1 **(b)**.

In order to measure the geometric deformation for this example, we compute the scalar curvature of $\mathcal{M} \subset \mathbb{R}^{256}$ and that of its three dimensional PCA projection $P(\mathcal{M})$. Figure 2 **(a)** shows the scalar curvature of the manifold $\mathcal{M}$ and the curvature of its projection $P(\mathcal{M})$ is displayed in Figure 2 **(b)**. Note that the scalar curvature of $\mathcal{M}$ shows some variations over its surface, but in overall a fairly constant (and positive) value, indicating a spherical geometry slightly deformed via the frequency modulation map $\mathcal{A}$. On the other hand, the PCA projection of $\mathcal{M}$ in a three dimensional space shows significant variations as expected from the graphical display in Figure 1 **(b)**. There are, in particular, two sets of four maximal scalar curvature values, corresponding to the corners of the cubic shaped surface shown in Figure 1 **(b)**.



(a) **(b)**

Figure 1: **(a)** The sphere $\mathbb{S}^2 \subset \mathbb{R}^3$; **(b)** The PCA projection $P(\mathcal{M}) \subset \mathbb{R}^3$ of the modulated sphere $\mathcal{M} \subset \mathbb{R}^{256}$.



(a) **(b)**

Figure 2: Modulation of the sphere $\mathbb{S}^2$. **(a)** The scalar curvature of $\mathcal{M} \subset \mathbb{R}^{256}$; **(b)** The scalar curvature of $P(\mathcal{M}) \subset \mathbb{R}^3$.

4.2 Frequency Modulation for a Torus

In this second numerical example, we use the surface of the torus $\Omega = \mathbb{T}^2 \subset \mathbb{R}^3$ in combination with the modulation map

$$\mathcal{A}_\alpha(t_i) = \sum_{k=1}^{3} \sin((\alpha_k^0 + \gamma\alpha_k)t_i + b_k),$$

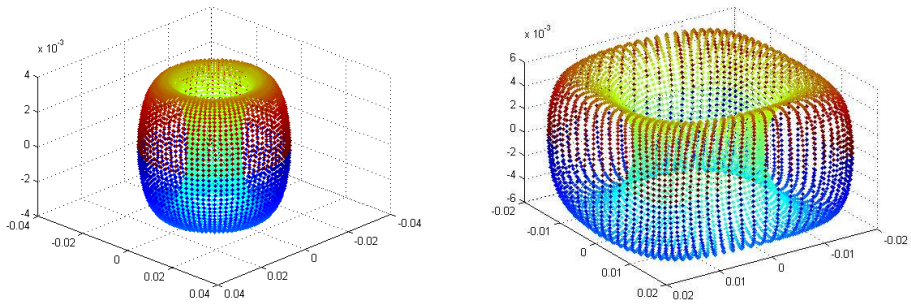
and the torus parametrization

$$\begin{aligned} \alpha_1(u,v) &= (R + r\cos(v))\cos(u), \\ \alpha_2(u,v) &= (R + r\cos(v))\sin(u), \\ \alpha_3(u,v) &= r\sin(v). \end{aligned}$$

Again, we use a finite and regular distribution of values $u \in [0, 2\pi]$, $v \in [0, 2\pi]$, and $\{t_i\}_{i=1}^{256} \subset [0, 1]$. As in the previous example, the parameter γ (the frequency bandwidth) plays a key role in the cubic deformation, as shown in Figure 3 **(b)**. The scalar curvature

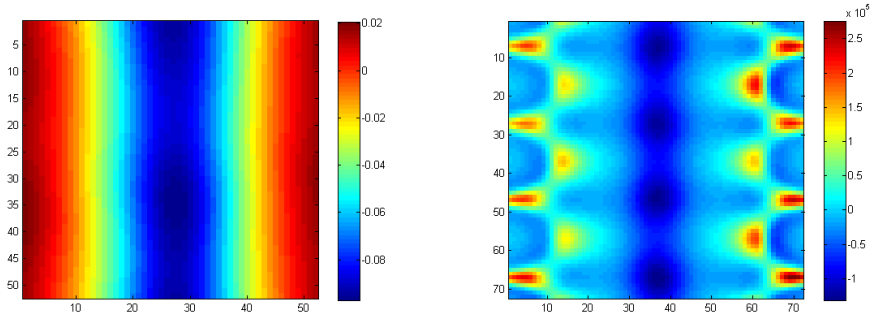
of $\mathcal{M} \subset \mathbb{R}^{256}$ is shown in Figure 4 **(a)**. Here, we observe a typical pattern for the torus geometry: a minimal (and negative) value for the scalar curvature corresponding to the smaller circle on the inside of the torus (depicted with the middle vertical line in Figure 4 **(a)**), two circles with zero curvature on the top and bottom sections of the torus (depicted with two vertical lines equidistant to the middle of Figure 4 **(a)**), and one circle with maximal (and positive) scalar curvature on the outside of the torus (depicted with the leftmost and rightmost vertical lines of the Figure 4 **(a)**).

The curvature of the PCA projection $P(\mathcal{M})$, shown in Figure 4 **(b)**, illustrates a similar structure, but with a considerable geometrical deformation which includes two sets of four points with maximal scalar curvature, representing the corners of the cubic shaped projection $P(\mathcal{M})$ shown in Figure 3 **(b)**.



(a) **(b)**

Figure 3: **(a)** The torus $\mathbb{T}^2 \subset \mathbb{R}^3$; **(b)** The PCA projection $P(\mathcal{M}) \subset \mathbb{R}^3$ of the modulated torus $\mathcal{M} \subset \mathbb{R}^{256}$.



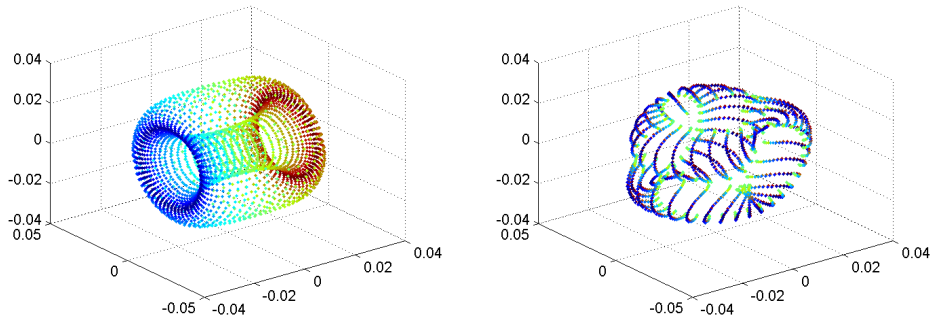
(a) **(b)**

Figure 4: Modulation of the torus $\mathbb{T}^2$. **(a)** The scalar curvature of $\mathcal{M} \subset \mathbb{R}^{256}$; **(b)** The scalar curvature of $P(\mathcal{M}) \subset \mathbb{R}^3$.

4.3 PCA and Isomap Projection of the Modulated Torus

We remark that a basic property of modulation manifolds is that an increase of the frequency bandwidth parameter γ amplifies the geometrical distortion of $\mathcal{M} \subset \mathbb{R}^n$. In

fact, it can be observed experimentally that for large bandwidth γ, standard projection methods, such as PCA, fail to recover the original geometry, unlike modern nonlinear dimensionality reduction methods, such as isomap. Figure 5 **(b)** shows how the linear PCA projection destroys the geometrical content of the modulated torus $\mathcal{M}$, whereas the nonlinear isomap projection achieves to recover (similar to the spirit of the classical Swiss Roll example) the topological and geometrical features of the torus $\mathbb{T}^2$ fairly well.



(a) **(b)**

Figure 5: Modulation of the torus $\mathbb{T}^2$. **(a)** Isomap projection of $\mathcal{M}$; **(b)** PCA projection of $\mathcal{M}$.

5 Conclusion and Future Steps

We have used frequency modulation maps to generate modulation manifolds and their relevant data sets for manifold learning. We have developed a numerical scheme, Algorithm 3.1, for computing the scalar curvature of a modulation manifold, along with its metric tensor. We applied the algorithm to two different test cases of frequency modulated manifolds. In one test case, we considered using the sphere $\mathbb{S}^2$, the other example relies on the torus $\mathbb{T}^2$. The numerical examples are illustrating the geometrical distortion incurred by the dimensionality reduction, when relying on PCA projections. Moreover, we have shown that the standard linear PCA projection is outperformed by the nonlinear somap projection, which achieves to recover the topological properties of the modulated surface, even at high frequency bandwidths, very well. In conclusion, the findings of this paper provide a first insight into the nature of frequency modulation maps and modulation manifolds. Further investigations along these lines will be made in future work. Note that the consideration of weaker structural assumptions on Ω and $\mathcal{M}$ is a very valuable but challenging task for a large variety of applications. In this context, the work on persistent homology [14] and discrete Morse theory [3] offers a suitable background that can be used in future steps.

6 Acknowledgments

The authors are supported by the priority program DFG-SPP 1324 of the Deutsche Forschungsgemeinschaft (DFG).

References

[1] D.S. Broomhead and M. Kirby. A new approach to dimensionality reduction: theory and algorithms. *SIAM Journal on Applied Mathematics*, 60(6):2114–2142, 2000.

[2] A. Brun, C.-F. Westin, M. Herberthson, and H. Knutsson. Fast manifold learning based on Riemannian normal coordinates. In *Proceedings of the SCIA;05*, pages 920–929, Joensuu, Finland, June 2005.

[3] R. Forman. A user's guide to discrete Morse theory. *Seminaire Lotharingien de Combinatoire*, 48:B48c, 2002.

[4] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, 2004.

[5] J.A. Lee and M. Verleysen. *Nonlinear Dimensionality Reduction*. Springer, 2007.

[6] J.M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Springer, 1997.

[7] T. Lin and H. Zha. Riemannian manifold learning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30:796–809, 2008.

[8] T. Lin, H. Zha, and S.U. Lee. Riemannian Manifold Learning for Nonlinear Dimensionality Reduction. *Lecture Notes in Computer Science*, 3951:44, 2006.

[9] P. Niyogi, S. Smale, and S. Weinberger. Finding the homology of submanifolds with high confidence from random samples. *Discrete and Computational Geometry*, 39(1):419–441, 2008.

[10] E. Saucan, E. Appleboim, and Y.Y. Zeevi. Geometric sampling of manifolds for image representation and processing. *Lecture notes in Computer Science*, 4485:907–918, 2007.

[11] E. Saucan, E. Appleboim, and Y.Y. Zeevi. Sampling and reconstruction of surfaces and higher dimensional manifolds. *Journal of Mathematical Imaging and Vision*, 30(1):105–123, 2008.

[12] G. Walschap. *Metric Structures in Differential Geometry*. Springer, 2004.

[13] H. Zha and Z. Zhang. Continuum Isomap for manifold learnings. *Computational Statistics and Data Analysis*, 52(1):184–200, 2007.

[14] A. Zomorodian and G. Carlsson. Computing persistent homology. *Discrete Comput. Geom.*, 33(2):249–274, 2005.