2007 International Symposium on Nonlinear Theory and its Applications
NOLTA'07, Vancouver, Canada, September 16-19, 2007

SOUND-MORPHING METHOD USING CONSTRUCTIVE MORSE THEORY

Yoshinao SHIRAKI

Department of Information Science, Shonan Institute of Technology.
1-1-25 Tsujido-Nishikaigan, Fujisawa, Kanagawa 251-8511, Japan
Email shiraki@info.shonan-it.ac.jp Tel: +81 466 30 0206, Fax: +81 466 34 5932

ABSTRACT

We investigate the relation between a speaker adaptation method such as VFS (Vector Field Smoothing [5]) and a sound-morphing method from a geometrical point of view. It is shown that topological field theory yields the systematic treatment of these two methods by the fact that the chain complex (Witten's complex) plays a central role in their deformations of spectra. Witten's complex combines the local and the global information of a manifold (or spectrum) algebraically. Consequently, it gives us numerical and algebraic invariants of the manifold. Some of the examples and the application to speech-spectra of classical mathematical ideas — for example, Morse theory and generalized Poincarè's conjecture in higher dimensions — for topological field theory are discussed.

1. INTRODUCTION

The object of this note is to construct a nonlinear map ∂ on a manifold and to show that a sequence of the map ∂ is a chain complex: $\partial \circ \partial = 0$.

A sound-morphing method (see Osaka [4] and Shiraki [8]) and a speaker adaptation scheme [5] both use essentially spectral deformation. One of the basic problems of these methods is how to deform speech-spectra gradually while maintaining intelligible speech as a specific individual. In this note, we consider the problem above to be the study of differentiable manifolds (spectra) and differentiable maps. Then, naturally, two manifolds are considered equivalent if they are diffeomorphic: there exists a differentiable map from one to the other with a differentiable inverse. So the problem turns out to be how to construct numerical and algebraic invariants determined by the diffeomorphism class of the manifold. We use Morse theory [3, 7] in order to extract topological invariants of each given speech-spectrum from their critical points, i.e., our focus is attention to the peaks and bottoms of speech-spectrum which are important characteristics of the perceptual property of speech. Morse theory relates collections of the local data around critical points and the global topological structure of the manifold. In this theory, Witten's complex plays a central role in constructing invariants of the manifolds [1, 11]. The construction of Witten's complex is numerical and algebraic, and the construction actually connects critical points of the manifolds. Therefore, Witten's complex gives constructive solutions to the problems of speaker adaptation and the sound-morphing mentioned above.

In Sect.2, we briefly review parts of Morse theory that are necessary preparations for describing deformations of speech-spectra, and show that a chain complex plays an essential role in connecting critical points. In Sect.3, we construct a nonlinear map (boundary operator) and show in Theorem 2 that a sequence of the operator is a chain complex.

2. PRELIMINARIES

In this section, let us briefly review the basic notions of algebraic topology and Morse theory [1, 2, 3, 6]. Let M be an n-dimensional compact smooth manifold, $\varphi \circ \psi$ be a composite, $f : M \to \mathbf{R}$ be a real-valued smooth map on an n-dimensional manifold M.

Definition 2.1: *A point $p \in M$ is said to be a* **critical point** *of f if, in some coordinate neighborhood $u = (u_1, \cdots, u_n)$ around p, $\frac{\partial f}{\partial u_i}(p) = 0 \ \ (i = 1, \cdots, n)$.*

Definition 2.2: *In a coordinate system u, a matrix $Hf_p = (\frac{\partial^2 f}{\partial u_i \partial u_j})$ is said to be the* **Hessian** *of f at a critical point p. The number $\mu(p)$ of negative-eigenvalue of Hessian Hf_p is said to be the* **Morse index** *of a critical point p. If Hf_p is nonsingular, a point p is said to be a* **nondegenerate** *critical point. A* **Morse function** *f on a smooth manifold M is a smooth function f such that the Hessian of f is nonsingular at every critical point.*

Definition 2.3: *A sequence as shown below of maps from one linear space to another linear is said to be a* **chain complex** *if $\partial_{k-1} \circ \partial_k = 0 (1 \leq k \leq n)$, denoted it by (C_*, ∂) or C_*. C_k is said to be a k-dimensional* **chain group** *and ∂ is said to be a* **boundary operator**.

$$C_n \stackrel{\partial_n}{\to} C_{n-1} \stackrel{\partial_{n-1}}{\to} \ldots \to C_k \stackrel{\partial_k}{\to} C_{k-1} \to \ldots \to C_1 \stackrel{\partial_1}{\to} C_0$$

2.1. Morse Theory and Witten's complex

Let M be an n-dimensional closed orientable manifold and consider a Morse function f on M. We consider the following ODE

$$\frac{d}{dt}\phi(t) = -\nabla_{\phi(t)} f, \tag{1}$$

where ϕ is a curve of $\phi : \ \to M$. is not compact, so it seems natural that we consider the boundary condition to construct a solution space of equation (1). For this purpose, we use critical points of the Morse function. A set of critical points is given by the following equation:

$$Cr(M, f) := \{p \in M | df(p) = 0\}. \tag{2}$$

Because the function f is a Morse function, the set $Cr(M,f)$ is finite. For $p \in Cr(M,f)$, let $\mu(p)$ be the Morse index of p. Namely, $\mu(p)$ is the number of negative-eigenvalue of Hessian. Using this as the boundary condition, we define $\mathcal{M}(M, f : p, q)$ as follows for two critical points $p, q \in Cr(M,f)$

$$\mathcal{M}(M, f : p, q) := \left\{ \; l : \; \rightarrow M \left| \begin{array}{l} \frac{dl}{dt} = -\nabla f \\ \lim\limits_{t \to -\infty} \phi(t) = p \\ \lim\limits_{t \to \infty} \phi(t) = q \end{array} \right. \right\}. \tag{3}$$

Now we count the cardinality of the set $\mathcal{M}(M, f : p, q)$ has an action on $\mathcal{M}(M, f : p, q)$ as $\phi(t+s) = \phi(t) \circ \phi(s)$, so we consider the quotient space $\bar{\mathcal{M}}(M, f : p, q)$ with respect to this action. In general, to count the cardinality of set, the dimension of the set must be 0-dimension. Because the set of critical points depends on the choice of a Morse function, the number of the solution space is not a topological invariant.

So we do algebraic construction. Let us define the graded Abelian group $C_\bullet(M; f)$ as follows:

$$C_k(M; f) := \bigoplus_{\substack{p \in Cr(M,f) \\ \mu(p) = k}} \cdot [p], \tag{4}$$

where the right-hand side is the free module on in which the number of generating elements is equal to the number of critical points of index k. We define a map $\partial_k : C_k(M; f) \rightarrow C_{k-1}(M; f)$ on the group as follows:

$$\partial_k([p]) := \sum_{\substack{q \in Cr(M,f) \\ \mu(q) = k-1}} \sharp \, \bar{\mathcal{M}}(M, f : p, q)[q], \tag{5}$$

where $\sharp$ means counting the number; more precisely speaking, consider $\bar{\mathcal{M}}(M, f : p, q)[p]$ to be a 0-dimensional oriented manifold and $\sharp$ means counting the number of signed number of the manifold.

The formulation above of chain complex is due to Witten[11]. In this construction of the chain complex, the following theorem holds.

Theorem 1. *$(C_\bullet(M; f), \partial)$ is a chain complex. Namely, $\partial_{k-1} \circ \partial_k = 0$. Furthermore, homology group of $(C_\bullet(M; f), \partial)$ is isomorphic to homology group of M: $H_*((C_\bullet(M; f), \partial)) \cong H_*(M)$.*

2.2. Stable and unstable manifold

Let us define a stable manifold S_p and an unstable manifold U_p for a critical point p of an n-dimensional manifold M as follows:

$$S_p = \left\{ x \in M \mid \lim_{t \to \infty} \phi(t) = p \right\}, \tag{6}$$

$$U_p = \left\{ x \in M \mid \lim_{t \to -\infty} \phi(t) = p \right\}. \tag{7}$$

When the Morse index of p is k, the dimensions of S_p and U_p are $n-k$ and k, respectively. These submanifolds of S_p and U_p give a certain kind of the decomposition of the manifold M and they provide us the systematic point of view when we study manifolds (see e.g. [10]). In fact, for a set Cr of critical points of a Morse function the following equation holds:

$$\cup_{p \in Cr} U_p = M. \tag{8}$$

This means that the unstable manifolds gives the cell (homeomorphic to $^{\mu}(p)$) decomposition of a manifold M.

Let $c \in$ be such that $f(q) < c < f(p)$, and denote a neighborhood $N := f^{-1}(c)$. Then S_q and U_p intersect N transversally (see e.g.[9]). We define $S_q(c) := N \cap S_q, U_p(c) := N \cap U_p$ and $\dim U_p(c) = k-1, \dim S_q(c) = n-k$. Therefore, if $S_q(c)$ and $U_p(c)$ intersect transversally, their intersection number is finite.

Morse-Smale type: *The gradient flow of a Morse function f : $M \rightarrow$ is said to be of Morse-Smale type if for any two critical points p and q the stable and unstable manifolds S_p and U_q intersect transversally.*

This requirement can be achieved by means of an arbitrarily small alternation of the Riemannian metric [10]. Hereafter we assume that a Morse function is of Morse-Smale type.

If two critical points p, q are with $\mu(p) - \mu(q) = i$ ($i = 1, 2, 3$), for the intersection of the stable and unstable manifold the following **Lemma 1**,**Lemma 2** and **Lemma 3** hold [1]:

Lemma 1:*For $\mu(p) = \mu(q)$, $U_p \cap S_q = \emptyset$. For $\mu(p) - \mu(q) = 1$, the following equations hold:*

$$U_p \cap S_q = \cup J_i(p, q), \; J_i(p, q) \cap J_j(p, q) = \emptyset \; (i \neq j),$$

where each of $J_i(p, q)$ is a flow trajectory which is a 1-dimensional submanifold of M and homeomorphic to .

Lemma 2: *For $\mu(p) - \mu(r) = 2$,*

$$U_p \cap S_r = \cup_l T(l), \; T(l) \cap T(l') = \emptyset \; (l \neq l'),$$

where each of $T(l)$ is a 2-dimensional submanifold of M and homeomorphic to ^{2}.

Now we fix a point q with $\mu(p) - \mu(q) = 1$. Then $\overline{J_i(p,q)} \cup \overline{J_j(q,r)}$ is said to be a *edge* of $T(l)$ if for any $t \in$ there exists the gradient flow $\phi_{q_\nu}(t)$ which converges to $\overline{J_i(p,q)} \cup \overline{J_j(q,r)}$. Here, a point sequence $\{q_\nu\}$ in $T(l)$ converges to q.

Lemma 3:*(1)Each of $T(l)$ has just two edges. (2)There exists just one $T(l)$ whose edge is $\overline{J_i(p,q)} \cup \overline{J_j(q,r)}$.*

3. CONSTRUCTION OF BOUNDARY OPERATOR

In this section, we try to construct a map ∂ on a manifold M following Morse theory and Witten's complex. We use Morse theory as the basis of the construction and also use a set of critical points of Morse function as the boundary conditions. Hereafter let a manifold M be orientable.

3.1. Witten's complex based on orientation

We set a Morse function f on a manifold M so that S_q and U_p intersect transversally for any critical points p, q. We define Cr_k as follows:

$$Cr_k := \{p \in Cr(M, f) | \mu(p) = k\}, \tag{9}$$

and let the linear space C_k whose basis are Cr_k be defined as follows:

$$C_k := \{ \sum_{p \in Cr_k} a_p [p] | a_p \in \}. \tag{10}$$

We define the boundary operator ∂ based on the two orientations defined on an image of the gradient flow $\phi(t)$ as follows:

Procedure of Witten's complex based on orientations

1. Determine the orientation of a manifold M.
2. For each critical point $p \in Cr$ determine arbitrarily the orientation of the unstable manifold U_p.
3. Determine the orientation of the stable manifold S_p, which was determined from the intersection of the orientations of S_p and U_p, at p is $+1$, because S_p intersects U_p transversally.
4. For $p, q \in Cr$ determine the orientation of $S_q \cap U_p$ so that this orientation is equal to that of the intersection of the orientations of S_q and U_p.
5. When $\mu(p) - \mu(q) = 1$ holds, by **Lemma 1** the following equation holds:
$$S_q \cap U_p = \cup J_i(p,q). \tag{11}$$
Therefore, the orientation determined on $S_q \cap U_p$ determines the orientation of each element $J_i(p,q)$. Note that $J_i(p,q)$ has its own orientation preserving the gradient flow $\phi(t)$.
6. Assign a number $+1$ or -1 to $\varepsilon(J_i(p,q))$ according to whether the two orientations of $J_i(p,q)$ coincide: one orientation of $J_i(p,q)$ is determined by *Step4* and another is determined by $J_i(p,q)$ itself.

When $\mu(p) - \mu(q) = 1$, using $\varepsilon(J_i(p,q))$ defined in *Step6* let us define
$$m(p,q) := \sum_i \varepsilon(J_i(p,q)). \tag{12}$$
We define the boundary operation $\partial : C_k \to C_{k-1}$ as follows:
$$\partial(\sum_{p \in Cr_k} a_p[p]) := \sum_{p \in Cr_k} a_p \partial[p], \tag{13}$$
$$\partial[p] := \sum_{q \in Cr_{k-1}} m(p,q)[q]. \tag{14}$$
In this construction of the complex, the following theorem holds.

Theorem 2. (C_*, ∂) *defined above is a chain complex. Namely,* $\partial_{k-1} \circ \partial_k = 0$.

Proof. For each critical point $p \in Cr_k$, the composite of boundary operation is as follows:
$$\partial \circ \partial p = \sum_{r \in Cr_{k-2}} (\sum_{q \in Cr_{k-1}} m(p,q)m(q,r))[r].$$
For two critical points p, r with $\mu(p) - \mu(r) = 2$, we just prove the following equation:
$$\sum_{q \in Cr_{k-1}} m(p,q)m(q,r) = 0.$$
Using equation (12), we can rewrite the equation above into the following form:
$$\sum_{q \in Cr_{k-1}} (\sum_{i,j} \varepsilon(J_i(p,q))\varepsilon(J_j(q,r))) = 0. \tag{15}$$

First step: The stable manifold S_r and unstable manifold U_p intersect transversally, so by **Lemma 1** and **Lemma 2** these intersections can be represented by 1-dimensional submanifolds J and 2-dimensional submanifolds $T(l)$.

For any two critical points with $p \in Cr_k$, r $\in Cr_{k-2}$, the stable manifold S_r and unstable manifold U_p intersect transversally at finite points because the Morse function is of Morse-Smale type. Consequently $S_r \cap U_p$ is represented by the sum of two 2-dimensional manifolds $T(l)$:
$$S_r \cap U_p = \cup_l T(l), \tag{16}$$
$$T(l) \cap T(l') = \emptyset \ \ (l \neq l'). \tag{17}$$
Now let us think of flow trajectories $J_i(p,q)$ contained in the closure $\overline{T(l)}$ of the 2-dimensional manifold $T(l)$. These trajectories connect two critical points p, q with $\mu(p) - \mu(q) = 1$ and at most two trajectories are contained in the closure $\overline{T(l)}$. It is also the same case that at most two trajectories $J_j(q,r)$ connecting two critical points q, r with $\mu(q) - \mu(r) = 1$ are contained in the closure $\overline{T(l)}$.

Second step: Next we count the number of 1-dimensional submanifolds J and 2-dimensional submanifolds $T(l)$ contained in equation (15).

Let the combination of two trajectories be $\overline{J_i(p,q)} \cup \overline{J_j(q,r)}$. Then by **Lemma 3** this combination is the edge of at least one 2-dimensional submanifold $T(l)$ and each submanifold $T(l)$ has just two edges, so let us represent these edges $\overline{J_i(p,q)} \cup \overline{J_j(q,r)}, \overline{J_{i'}(p,q')} \cup \overline{J_{j'}(q',r)}$. For these two edges we investigate the sign of four trajectories: $\varepsilon(J_i(p,q)),\ \varepsilon(J_j(q,r)),\ \varepsilon(J_{i'}(p,q')),\ \varepsilon(J_{j'}(q',r))$. Concerning the orientation of the unstable and stable manifolds, we first decide the orientations of each unstable manifold in Step2 and then the orientations of each stable manifold are determined in Step3. Consequently at the point q the orientation of manifold M is determined by the orientations of the stable manifold S_q and the unstable manifold U_q.

Third step: Check two orientations determined on the flow trajectory $J_i(p,q)$:one orientation is determined by the stable and unstable manifolds (Step5), the other is by the flow trajectory itself.

By intersecting the three orientations (of S_q, U_q, and M) decided above and the orientation of the stable manifold S_r, at the point q the orientation determined by S_q, $J_i(q,r)$ and the orientation of S_r are equal. And by intersecting the orientations of these and U_p, at the point q the orientation of $T(l)$ and the orientation determined by the two flow trajectories $J_i(p.q), J_j(q,r)$ are equal. This is the same for the orientation of $J_{i'}(p,q'), J_{j'}(q',r)$.

Since the sign $\varepsilon(J_i(p,q))$ is determined by the orientations of $J_i(p,q)$ decided above and the gradient flow of $J_i(p,q)$, the following equation holds:
$$\varepsilon(J_i(p,q))\varepsilon(J_j(q,r)) + \varepsilon(J_{i'}(p,q'))\varepsilon(J_{j'}(q',r)) = 0 \tag{18}$$

Since each term of equation (18) depends on just one $T(l)$ by **Lemma 3**, we obtain (15) by summing up equation (18) concerning $T(l)$. This completes the proof of Theorem 2.

□

3.2. Examples

The followings are some examples of Witten's complex for 1- and 2-manifold. We omit the case of a manifold with boundary.

3.2.1. 1-dimensional Case

Let M^1 be a 1-dimensional speech-spectrum envelope in log-scale. After normalizing M^1 properly, let f be a height function of M^1. Then, f is a Morse function. In fact, the Hessians at peak and bottom points are positive or negative definite, so all critical points are nondegenerate and the Morse index is 0 or 1.

In this case, there is a 1-dimensional homology group $H_1(M^1)$ corresponding to peaks, and each element of the homology is not independent. That is, the homology group is not changed if the number of peaks (or bottom) is increased or decreased by deformations of spectrum. This fact means that the homology group is homotopy equivalent and also holds in the case of 0 dimensional homology group corresponding to bottoms.

Let us consider the case of M^1 having two peaks. We denote two bottoms and peaks $p_1^0, p_2^0, p_1^1, p_2^1$. In this case each Morse index is 0, 0, 1, 1 and the calculation of Witten's complex is as follows:

$$\partial : C_1 \to C_0, \quad \partial p_1^1 = (+1)[p]_1^0 + (-1)[p]_2^0 = [p]_1^0 - [p]_2^0, \quad (19)$$
$$\partial p_2^1 = (+1)[p]_2^0 + (-1)[p]_1^0 = [p]_2^0 - [p]_1^0. \quad (20)$$

3.2.2. 2-dimensional Case

Let M^2 be a 2-dimensional power spectrum in log-scale. After normalizing M^2 properly, let f be a height function of M^2. Then f is a Morse function. In fact, the Hessians at peak and bottom points are positive or negative definite, so all critical points are nondegenerate and the Morse index is 0 or 1 or 2.

Let us consider the case of M^2 having two peaks. We denote two bottoms and peaks p^0, p^1, p_1^2, p_2^2. In this case each Morse index is 0, 1, 2, 2 and the calculation of Witten's complex is as follows:

$$\partial : C_2 \to C_1, \quad \partial p_1^2 = (-1)[p]^1 = -[p]^1, \quad \partial p_2^2 = (+1)[p]^1 = [p]^1, \quad (21)$$
$$\partial : C_1 \to C_0, \quad \partial p^1 = (+1 + (-1))[p]^0 = 0. \quad (22)$$

References

[1] A.Floer, "Witten's complex and infinite dimensional Morse theory," *J. Diff. Geom.*, **30**, pp.207-221, 1989

[2] K.Fukaya,"Floer homology of connected sum of homology 3-spheres," *Topology*, **Vol.35, No.1** , pp.89-136, 1996.01

[3] J.W.Milnor, *Morse Theory*, Ann. of Math. Studies 51, Princeton University Press, 1963

[4] N.Osaka, "A timbre interpolation using a sinusoidal model, " *IPSJ, SIGMUS*, 95-MUS-13, pp.45-50, 1995.12

[5] K.Ohkura, M.Sugiyama, and S.Sagayama, "Speaker Adaptation based on transfer vector field smoothing with continuous mixture density HMMs," in *Proc. of Int. Conf. Spoken Language Processing*, pp.369-372, 1992

[6] D.Salamon, "Morse theory, the Conley index and Floer homology," *Bull. London Math. Soc.***22**, pp.113-140, 1990

[7] Y.Shiraki,"Homotopy equivalent spectral transformation and Morse theory" in *IEICE Trans. Fundamentals* **Vol. E78-A, No.9**, pp.1186-1191, 1995.09

[8] Y.Shiraki,"Optimal Temporal Decomposition for Voice Morphing preserving Δ cepstrum" in *IEICE Trans. Fundamentals* **Vol. E87-A, No.3**, pp.577-583, 2004.03

[9] S.Smale, "Morse inequalities for a dynamical system," *Bull. Amer. Math. Soc.*, **66**, pp.43-49, 1960

[10] S.Smale, "On gradient dynamical systems," *Ann. Math.* **74**, pp.199-206, 1961

[11] E.Witten, "Supersymmetry and Morse theory," *J. Diff. Geom.*, **17**, pp.661-692, 1982