Invariants of Morse complexes, persistent homology and data analysis
Serguei Barannikov

To cite this version:
Serguei Barannikov. Invariants of Morse complexes, persistent homology and data analysis. Colloquium of the Steklov Mathematical Institute of Russian Academy of Sciences, Mar 2019, Moscow, Russia. 10.13140/RG.2.2.14363.31528. hal-02062497

HAL Id: hal-02062497
https://hal.science/hal-02062497
Submitted on 21 Mar 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
Invariants of Morse complexes, persistent homology and data analysis
Steklov Mathematics Institute

S.Barannikov (Paris VII, NRU HSE Moscow)

7 March 2019
Partition of critical values into pairs “birth-death” plus homological critical values (“births” paired with $+\infty$)
Partition of critical values into pairs “birth-death” plus homological critical values (“births” paired with $+\infty$)
The decomposition arises from bringing of the Morse complex over field $F$, defined by gradient trajectories of the function, to what I called “canonical form” by a linear transform respecting the filtration, given by order of the critical values.
“Canonical form” invariants of filtered complexes

- The decomposition arises from bringing of the Morse complex over field $F$, defined by gradient trajectories of the function, to what I called “canonical form” by a linear transform respecting the filtration, given by order of the critical values.

- These “canonical forms” are combinatorial invariants of filtered complexes.
Chain complexes

Recall: chain complex \((C_\ast, \partial_\ast)\) is a sequence of vector spaces over field \(F\) and linear operators

\[
C_{j+1} \xrightarrow{\partial_{j+1}} C_j \xrightarrow{\partial_j} C_{j-1} \rightarrow \ldots \rightarrow C_0
\]

which satisfy

\[
\partial_j \circ \partial_{j-1} = 0
\]

i.e. the image of \(\partial_{j+1}\) is contained in the kernel of \(\partial_j\):

\[
\text{im} (\partial_{j+1}) \subseteq \ker (\partial_j).
\]

The \(j\)-th homology of the chain complex \((C_\ast, \partial_\ast)\) is the quotient

\[
H_j = \ker (\partial_j) / \text{im} (\partial_{j+1}).
\]
Morse complex

\( f : M^n \rightarrow \mathbb{R} \), \( f \in C^\infty \), generic, \( \{ x \mid f(x) \leq c \} \) — compact. Then \( p_\alpha \)-critical points, \( df \mid_{T_{p_\alpha}} = 0 \), are isolated, near.

\( p_\alpha : f = \sum_{i=1}^{j} -(x^i)^2 + \sum_{i=j}^{n} (x^i)^2 \). Let \( g \) is a generic metric. Then

\[
C_j = \bigoplus_{\text{index}(p_\alpha) = j} [p_\alpha, \text{or}(T_{p_\alpha}^-)]
\]

where \( T_{p_\alpha} = T_{p_\alpha}^- \oplus T_{p_\alpha}^+ \) is wrt \( \partial^2 f \) and \( g \).

The differential is

\[
\partial_j [p_\alpha, \text{or}] = \sum_{\text{index}(p_\beta) = j-1} [p_\beta, \text{or}] \# \mathcal{M}(p_\alpha, p_\beta)
\]

\[
\mathcal{M}(p_\alpha, p_\beta) = \{ \gamma : \mathbb{R} \rightarrow M^n \mid \\
\dot{\gamma} = -(\text{grad}_g f)(\gamma(t)), \lim_{t \to -\infty} = p_\alpha, \lim_{t \to +\infty} = p_\beta \} / \mathbb{R}
\]
Let $C_\ast$ be a filtered chain complex, an increasing sequence of subcomplexes $F_s C_\ast \subset F_r C_\ast$, $s < r$, indexed by a finite set of real numbers, $F_{\text{max}} C_\ast = C_\ast$. It can come with a basis compatible with filtration so that each subspace $F_r C_j$ is the span $\langle e_1^{(j)}, \ldots, e_{i_r}^{(j)} \rangle$. 

Theorem (SB, 1994) There is a basis compatible with filtration in which the complex takes the "canonical form".
"Canonical form" invariants of filtered complexes [SB1994]

Let $C_\ast$ be a filtered chain complex, an increasing sequence of subcomplexes $F_s C_\ast \subset F_r C_\ast$, $s < r$, indexed by a finite set of real numbers, $F_{\text{max}} C_\ast = C_\ast$. It can come with a basis compatible with filtration so that each subspace $F_r C_j$ is the span $\left\langle e_1^{(j)}, \ldots, e_i^{(j)} \right\rangle$

A chain complex with bases $\{\tilde{e}_i^{(j)}\}_{i \in \{1, \ldots, \text{dim}_F C_j\}}$ is in "canonical form" if for any basis element $\tilde{e}_i^{(j)}$ either $\partial \tilde{e}_i^{(j)} = 0$ or $\partial \tilde{e}_i^{(j)} = \tilde{e}_{i,1}^{(j-1)}$. In the latter case, $\tilde{e}_i^{(j)} \neq \tilde{e}_{i,1}^{(j)} \Rightarrow \partial \tilde{e}_i^{(j)} \neq \partial \tilde{e}_{i,1}^{(j)}$. Theorem (SB, 1994) There is a basis compatible with filtration in which the complex takes the "canonical form". This "canonical form", i.e. the pairing "birth-death" between indices of filtration plus filtration indices of homology generators, is uniquely determined.
“Canonical form” invariants of filtered complexes [SB1994]

- Let $C_\ast$ be a filtered chain complex, an increasing sequence of subcomplexes $F_s C_\ast \subset F_r C_\ast$, $s < r$, indexed by a finite set of real numbers, $F_{\text{max}} C_\ast = C_\ast$. It can come with a basis compatible with filtration so that each subspace $F_r C_j$ is the span $\langle e_1^{(j)}, \ldots, e_{i_r}^{(j)} \rangle$

- Chain complex with bases $\{ \tilde{e}_i^{(j)} \}_{i \in \{ 0, 1, \ldots, \dim F C_j \}}$ is in “canonical form” if for any basis element $\tilde{e}_i^{(j)}$ either $\partial \tilde{e}_i^{(j)} = 0$ or $\partial \tilde{e}_i^{(j)} = \tilde{e}_i^{(j-1)}$. In the latter case, $\tilde{e}_i^{(j)} \neq \tilde{e}_{i_1}^{(j)} \Rightarrow \partial \tilde{e}_i^{(j)} \neq \partial \tilde{e}_{i_1}^{(j)}$.

- Theorem (SB, 1994) There is a basis compatible with filtration in which the complex takes the “canonical form”. This “canonical form”, i.e. the pairing “birth-death” between indices of filtration plus filtration indices of homology generators, is uniquely determined.
There are three equivalent visualizations of the same invariants. "Persistence Bar-codes"/"Persistence diagrams" were introduced in the beginning of 2000s (H. Edelsbrunner, J. Harer, A. Zamorodian “Hierarchical Morse complexes for piecewise linear 2-manifolds” Proc. of Symp on Comput Geometry, June 2001). There are several software packages for computing these invariants of a finite filtration. The principal algorithm is based on the bringing of the filtered complex to its canonical form by upper-triangular matrices from [SB1994].
“Canonical form” invariants of Morse complexes

- The Morse complex is naturally filtered by the set \( \{f(p_\alpha)\} \) of critical values of \( f \): \([p_\alpha, \text{or}] \in F_s C_* \text{ if } f(p_\alpha) \leq s\)
The Morse complex is naturally filtered by the set \( \{ f(p_\alpha) \} \) of critical values of \( f: [p_\alpha, \infty] \in F_s C_* \) if \( f(p_\alpha) \leq s \).

- Canonical partition of the set of critical values \( \{ f(p_\alpha) \} \) into pairs “birth-death”, plus separate set giving a basis \( H(M, F) \) or “births” paired with \( +\infty \).

Claim: the “canonical form” of Morse complex does not depend on the metrics: under generic perturbation of the metrics the complex changes via series of change of bases.
The Morse complex is naturally filtered by the set \( \{ f(p_{\alpha}) \} \) of critical values of \( f: \) \([p_{\alpha}, \infty] \in F_s C_* \text{ if } f(p_{\alpha}) \leq s \)

→ canonical partition of the set of critical values \( \{ f(p_{\alpha}) \} \) into pairs “birth-death”, plus separate set giving a basis \( H(M, F) \) or “births” paired with \( +\infty \)

Claim: the “canonical form” of Morse complex does not depend on the metrics: under generic perturbation of the metrics the complex changes via series of change of bases: \( e_i^{(j)} \rightarrow e_i^{(j)} + e_{\text{lower}}^{(j)} \)
“Canonical form” invariants of Morse complexes

The Morse complex is naturally filtered by the set \( \{ f(p_\alpha) \} \) of critical values of \( f: [p_\alpha, \infty] \in F_s C_* \) if \( f(p_\alpha) \leq s \).

- Canonical partition of the set of critical values \( \{ f(p_\alpha) \} \) into pairs “birth-death”, plus separate set giving a basis \( H(M, F) \) or “births” paired with \( +\infty \).

Claim: the “canonical form” of Morse complex does not depend on the metrics: under generic perturbation of the metrics the complex changes via series of change of bases: 
\[
E_i^{(j)} \to E_i^{(j)} + E_{i_{\text{lower}}}^{(j)}
\]

when the function is deformed the “canonical form” invariant changes naturally in continuous way. This can be expressed in \( \epsilon - \delta \) language.
Point clouds and Čech Complex

- Point cloud $\rightarrow$ filtered simplicial complex, $f = \text{distance}$
Point clouds and Čech Complex

- Point cloud → filtered simplicial complex, $f = \text{distance}$

- More generally, simplicial complex from collection of subsets of a discrete set, closed under restrictions, equipped with distances between points of the set.
Arnold’s problem on extension of smooth function inside a ball

Given \( f \in C^\infty (\partial B^n \times [-\varepsilon, \varepsilon]) \) → how many are there critical points of generic smooth extension of \( f \) inside the ball \( B \)?
Arnold’s problem on extension of smooth function inside a ball

- Given $f \in C^\infty(\partial B^n \times [-\varepsilon, \varepsilon]) \rightarrow$ how many are there critical points of generic smooth extension of $f$ inside the ball $B$?
- Theorem (SB, 1994) Pairs in “canonical form” $f |_{\partial B^n}$ indicating critical points in $B^n$:

They can cancell each other in certain configurations, details are in [SB1994].
Small eigenvalues of twisted Laplacian

These “canonical form” invariants were applied in Le Peutrec D., Nier F., Viterbo C. “The Witten Laplacian and Morse–Barannikov Complex” [LNV2011] to find formulas for small eigenvalues of the Witten Laplacian of \( d_{f,h} = hd + df \)

\[
\Delta_{f,h} = (d_{f,h} + d_{f,h}^*)^2 = d_{f,h}^*d_{f,h} + d_{f,h}d_{f,h}^* = \bigoplus_{p=0}^{\dim M} \Delta_{f,h}^{(p)}.
\]

There is a one to one correspondance \( j_p \) between \( \mathcal{U}^{(p)} \) and the set of eigenvalues (counted with multiplicities) of \( \Delta_{f,h}^{(p)} \) lying in \([0, h^{3/2}]\) such that

\[
j_p(U^{(p)}) = 0 \quad \text{if} \quad U^{(p)} \in \mathcal{U}_{H}^{(p)}
\]
\[
j_p(U^{(p)}) = \kappa^2(U^{(p+1)}) \frac{h}{\pi} \frac{|\lambda_1^{(p+1)} \ldots \lambda_{p+1}^{(p+1)}|}{|\lambda_1^{(p)} \ldots \lambda_p^{(p)}|} \frac{|\text{Hess}_f(U^{(p)})|^{1/2}}{|\text{Hess}_f(U^{(p+1)})|^{1/2}} (1 + \mathcal{O}(h)) e^{-2f(U^{(p+1)} - f(U^{(p)})} \quad \text{if} \quad \partial_B U^{(p+1)} = U^{(p)}
\]
\[
j_p(U^{(p)}) = \kappa^2(U^{(p)}) \frac{h}{\pi} \frac{|\lambda_1^{(p)} \ldots \lambda_p^{(p)}|}{|\lambda_1^{(p-1)} \ldots \lambda_{p-1}^{(p-1)}|} \frac{|\text{Hess}_f(U^{(p-1)})|^{1/2}}{|\text{Hess}_f(U^{(p)})|^{1/2}} (1 + \mathcal{O}(h)) e^{-2f(U^{(p)} - f(U^{(p-1)})} \quad \text{if} \quad \partial_B U^{(p)} = U^{(p-1)}
\]

Here the \( \lambda \)'s denote the negative eigenvalues of the Hess\( f \) at the corresponding points.
Arnold’s 4 cusps conjecture

conjecture solved in [ChP], Arnold described it in ([Arnold2002], page 79):

для любого однопараметрического гладкого семейства лежандровых кривых (интегральных кривых естественной контактной структуры в пространстве контактных элементов), связывающих лежандрову кривую фронта, движущуюся на плоскости внутри диска, ограниченного ею, с лежандровой кривой вывернутого фронта, движущейся наружу, фронты некоторых промежуточных лежандровых кривых из связывающего семейства имеют по крайней мере четыре точки возврата каждый.

Доказательство этой замечательной, имеющей глубокий физический смысл топологической теоремы очень сложно и использует, с одной стороны, недавний прогресс в симплектической топологии (гомологии Флёра и квантовые гомологии и т. п.), а с другой — результаты С. Баранникова об алгебре комплекса Морса (не получившие, к сожалению, заслуженного признания в момент своего появления несколько лет назад).
Klein bottle in data set of high contrast natural images

& al data set: $8 \cdot 10^6$ points in $\mathbb{R}^9$: ‘$3 \times 3$’ patches of high contrast in $4 \cdot 10^3$ natural images. Normalized to mean intensity and projected to $S^7$. First using these invariants on the most dense part of data set a circle of higher density is identified, then two more intersecting circles and finally a Klein bottle [CISZ]
Persistent cosmic web (cosmology, Sousbie & al[SPK])

Figure 3. An ascending 2-manifold (i.e. blue 2D wall) and an ascending 3-manifold (i.e. green 3D void) identified in a 512³ particles $100 \, h^{-1} \text{ Mpc}$ $\Lambda$CDM dark matter simulation. The manifolds where computed from a 64³ particles sub-sample.
Persistent cosmic web (cosmology, Sousbie & al)

Figure 10. The detected filamentary structure at a significance level of 5σ and three voids within a portion of SDSS DR7. Note that only the upper half of the distribution shown on figure 7 is displayed here for clarity reasons. The color of the filaments corresponds to the logarithm of the density field. The filaments of the SDSS extracted with DiaPerSE is readily available online at the URL http://www.tlp.fr/users/sousbie/.


