HAL
open science

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.

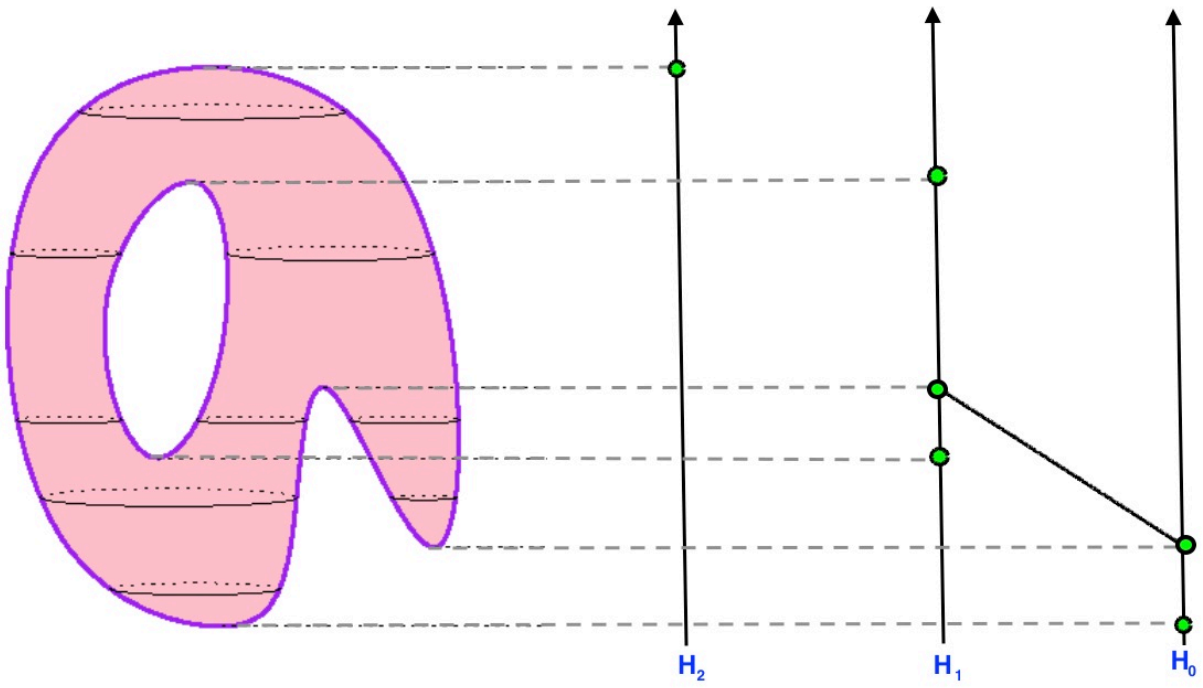
Invariants of Morse complexes, persistent homology and data analysis

Steklov Mathematics Institute

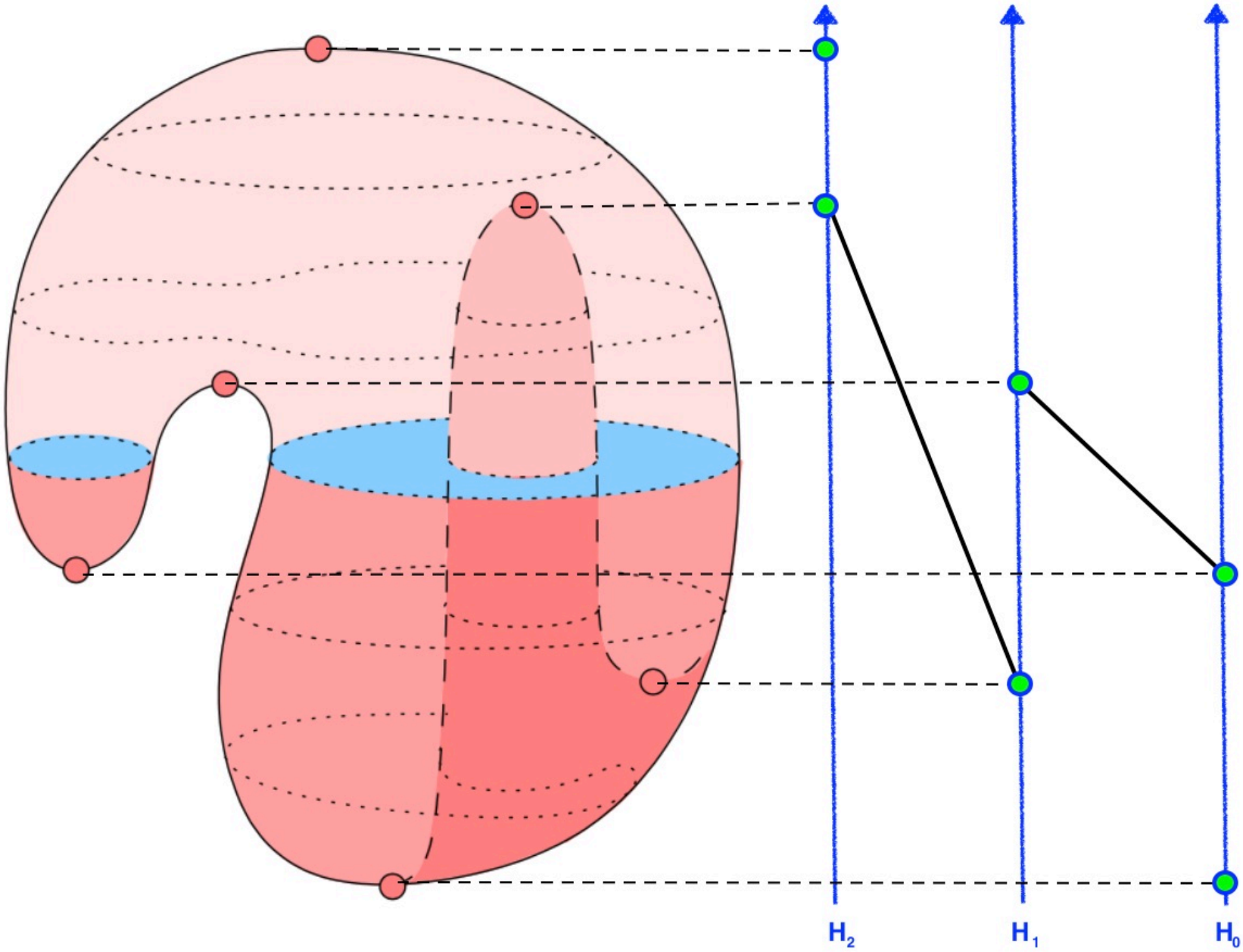
S.Barannikov (Paris VII, NRU HSE Moscow)

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 $+\infty$)



"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.

"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_*, ∂_*) is a sequence of vector spaces over field$F$ and linear operators

$$\to C_{j+1} \stackrel{\partial_{j+1}}{\to} C_j \stackrel{\partial_j}{\to} C_{j-1} \to \ldots \to C_0$$

which satisfy

$$\partial_j \circ \partial_{j-1} = 0$$

i.e. the image of ∂_{j+1} is contained in the kernel of ∂_j: $\mathrm{im}\,(\partial_{j+1}) \subseteq \ker\,(\partial_j)$. The $j-$th homology of the chain complex (C_*, ∂_*) is the quotient

$$H_j = \ker\,(\partial_j)\,/\,\mathrm{im}\,(\partial_{j+1})\,.$$

Morse complex

$f : M^n \to \mathbb{R}$, $f \in C^\infty$, generic, $\{x \mid f(x) \leq c\}$ −compact. Then p_α−critical points, $df \mid_{T_{p_\alpha}} = 0$, are isolated, near p_α:$f = \sum_{l=1}^{j} -(x^l)^2 + \sum_{l=j}^{n} (x^l)^2$. Let g is a generic metric. Then

$$C_j = \oplus_{\mathrm{index}(p_\alpha)=j} \left[p_\alpha, \mathrm{or}(T^-_{p_\alpha})\right]$$

where $T_{p_\alpha} = T^-_{p_\alpha} \oplus T^+_{p_\alpha}$ is wrt $\partial^2 f$ and g.
The differential is

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

$$\mathcal{M}(p_\alpha, p_\beta) = \{\gamma : \mathbb{R} \to M^n \mid$$
$$\dot{\gamma} = -(\mathrm{grad}_g f)(\gamma(t)), \lim_{t\to-\infty} = p_\alpha, \lim_{t\to+\infty} = p_\beta \Big\} / \mathbb{R}$$

"Canonical form" invariants of filtered complexes [SB1994]

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

"Canonical form" invariants of filtered complexes [SB1994]

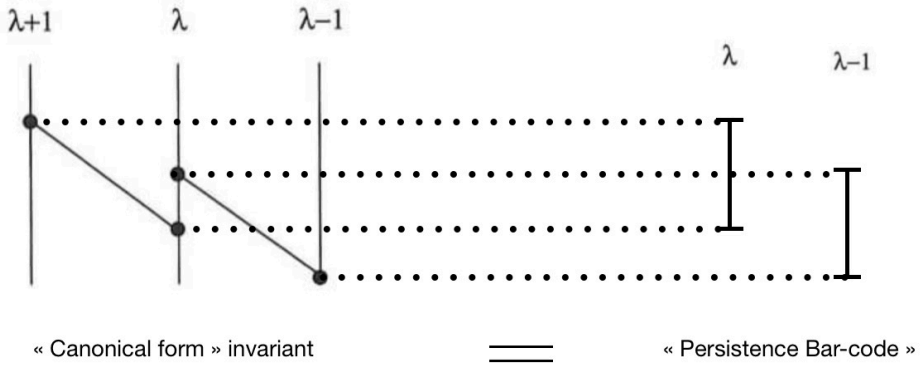
- Let C_* is a filtered chain complex,- an increasing sequence of subcomplexes $F_s C_* \subset F_r C_*$, $s < r$, indexed by finite set of real numbers, $F_{max} C_* = C_*$. It can come with a basis compatible with filtration so that each subspace $F_r C_j$ is the span $\left\langle e_1^{(j)}, \ldots, e_{i_r}^{(j)} \right\rangle$
- Chain complex with bases $\{\tilde{e}_i^{(j)}\}_{i \in \{0,1,\ldots\}}^{i \in 1,\ldots,\dim_F C_j}$ 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, $\tilde{e}_i^{(j)} \neq \tilde{e}_{i_1}^{(j)} \Rightarrow \partial \tilde{e}_i^{(j)} \neq \partial \tilde{e}_{i_1}^{(j)}$.

"Canonical form" invariants of filtered complexes [SB1994]

- Let C_* is a filtered chain complex,- an increasing sequence of subcomplexes $F_s C_* \subset F_r C_*$, $s < r$, indexed by finite set of real numbers, $F_{max} C_* = C_*$. It can come with a basis compatible with filtration so that each subspace $F_r C_j$ is the span $\left\langle e_1^{(j)}, \ldots, e_{i_r}^{(j)} \right\rangle$
- Chain complex with bases $\{\tilde{e}_i^{(j)}\}_{i \in \{0,1,\ldots\}}^{i \in 1,\ldots,\dim_F C_j}$ 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**.

"Canonical form" invariants ="Persistence Bar-codes"

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: $[p_\alpha, \mathrm{or}] \in F_s C_*$ if $f(p_\alpha) \leq s$

"Canonical form" invariants of Morse complexes

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

"Canonical form" invariants of Morse complexes

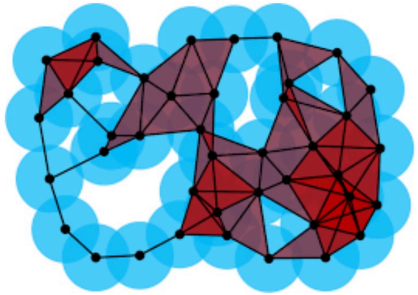
- The Morse complex is naturally filtered by the set $\{f(p_\alpha)\}$ of critical values of f: $[p_\alpha, \mathrm{or}] \in F_s C_*$if $f(p_\alpha) \leq s$
- $\rightarrow$canonical partition of the set of critical values $\{f(p_\alpha)\}$ into pairs "birth-death", plus separate set giving a basis $H(M, F)$ 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 chage of bases:$e_i^{(j)} \rightarrow e_i^{(j)} + e_{l_{\mathrm{lower}}}^{(j)}$

"Canonical form" invariants of Morse complexes

- The Morse complex is naturally filtered by the set $\{f(p_\alpha)\}$ of critical values of f: $[p_\alpha, \mathrm{or}] \in F_s C_*$if $f(p_\alpha) \leq s$
- $\rightarrow$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 chage of bases:$e_i^{(j)} \rightarrow e_i^{(j)} + e_{l_{\mathrm{lower}}}^{(j)}$
- when the function is deformed the "canonical form" invariant changes naturally in continuous way. This can be expressed in $\epsilon - \delta$ language.

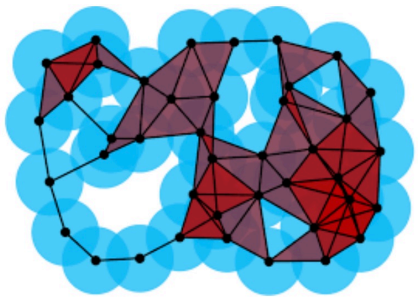
Point clouds and Čech Complex

- Point cloud $\rightarrow$filtered simplicial complex, $f = \mathrm{distance}$



Point clouds and Čech Complex

- Point cloud $\rightarrow$filtered simplicial complex, $f = \mathrm{distance}$

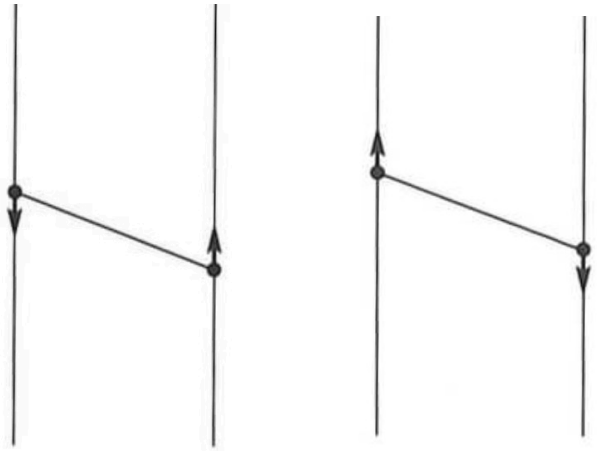


- More generally, simplicial complex from collection of subsets of a discrete set, closed under restrictions, equipped with distances between points of the set.

Arnold's problem on extension of smooth function inside a ball

- Given $f \in C^{\infty}\left(\partial B^n \times [-\varepsilon, \varepsilon]\right) \rightarrow$how many are there crtitical 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 crtitical points of generic smooth extension of f inside the ball B?
- Theorem (SB,1994) Pairs in "canonical form" $f\,|_{\partial B^n}$ indicating crititical 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 $d_{f,h} = hd + df$

$$\Delta_{f,h} = (d_{f,h} + d^*_{f,h})^2 = d^*_{f,h} d_{f,h} + d_{f,h} d^*_{f,h} = \bigoplus_{p=0}^{\dim M} \Delta^{(p)}_{f,h}.$$

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

$$j_p(U^{(p)}) = 0 \quad \text{if} \quad U^{(p)} \in \mathcal{U}^{(p)}_H$$

$$j_p(U^{(p)}) = \kappa^2(U^{(p+1)}) \frac{h}{\pi} \frac{|\lambda^{(p+1)}_1 \dots \lambda^{(p+1)}_{p+1}|}{|\lambda^{(p)}_1 \dots \lambda^{(p)}_p|} \frac{|\mathrm{Hess} f(U^{(p)})|^{1/2}}{|\mathrm{Hess} f(U^{(p+1)})|^{1/2}} (1 + \mathcal{O}(h)) e^{-2\frac{f(U^{(p+1)}) - f(U^{(p)})}{h}}$$

$$\text{if } \partial_B U^{(p+1)} = U^{(p)}$$

$$j_p(U^{(p)}) = \kappa^2(U^{(p)}) \frac{h}{\pi} \frac{|\lambda^{(p)}_1 \dots \lambda^{(p)}_p|}{|\lambda^{(p-1)}_1 \dots \lambda^{(p-1)}_{p-1}|} \frac{|\mathrm{Hess} f(U^{(p-1)})|^{1/2}}{|\mathrm{Hess} f(U^{(p)})|^{1/2}} (1 + \mathcal{O}(h)) e^{-2\frac{f(U^{(p)}) - f(U^{(p-1)})}{h}}$$

$$\text{if } \partial_B U^{(p)} = U^{(p-1)}$$

Here the λ's denote the negative eigenvalues of the $\mathrm{Hess} f$ 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 внутрь 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

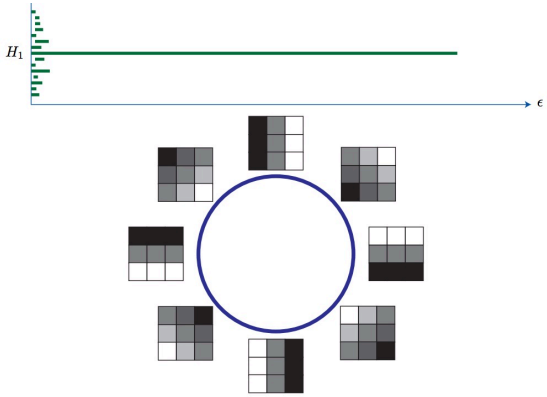


FIGURE 5. The H_1 barcode for a random sampling of 5000 points of $\mathcal{M}[300, 25]$ yields a single generator. This generator indicates the nodal line between a single light and single dark patch as being the dominant feature of the primary circle in $\mathcal{M}$.

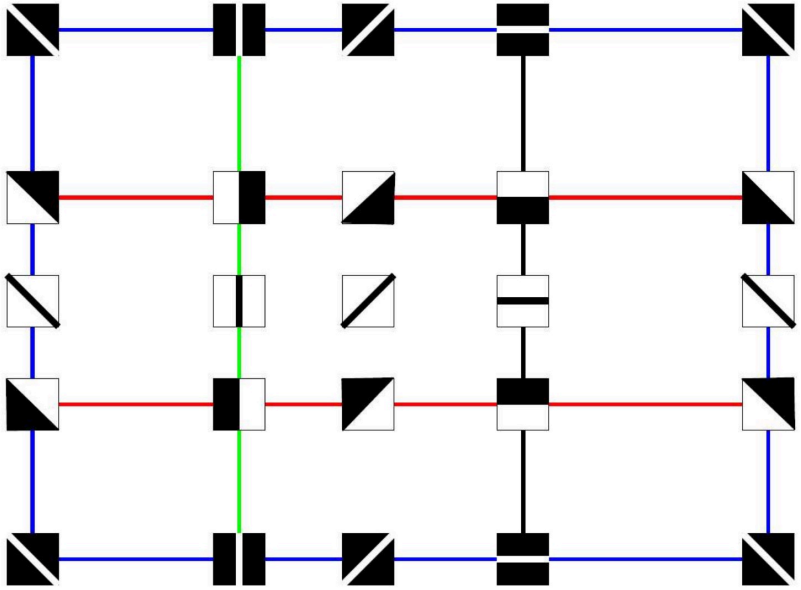
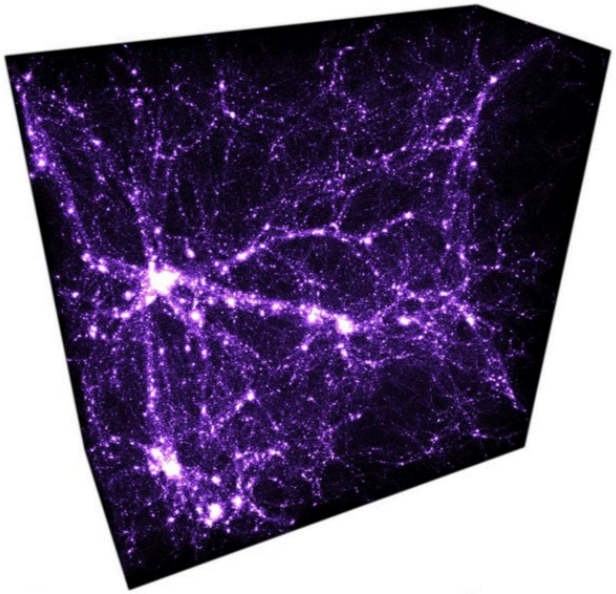


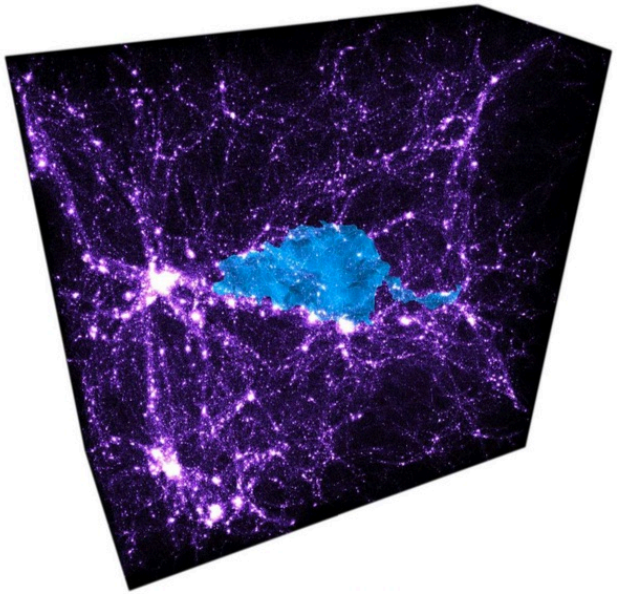
Figure 6: 3 by 3 patches parametrized by the Klein bottle

Mumford & al data set: $8 \cdot 10^6$ points in $\mathbb{R}^9$: '3×3' 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]

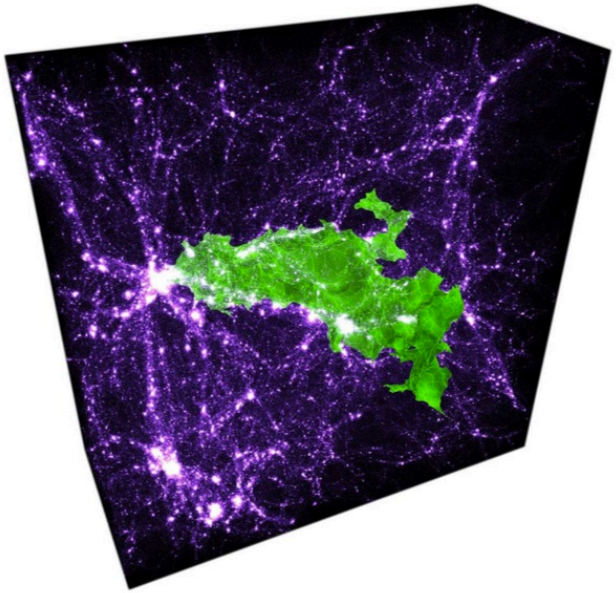
Persistent cosmic web (cosmology, Sousbie & al[SPK])



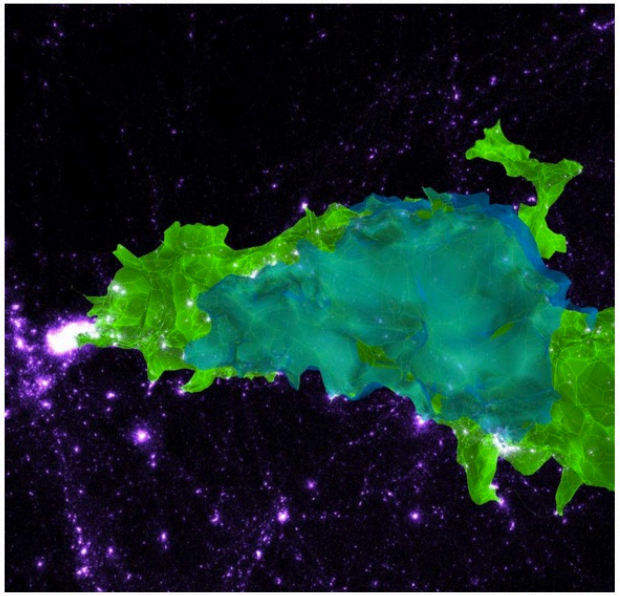
(a) Dark matter distribution in a $50 \times 50 \times 20\, h^{-1}$ Mpc sub-box



(b) An ascending 2-manifold (*i.e.* a wall)



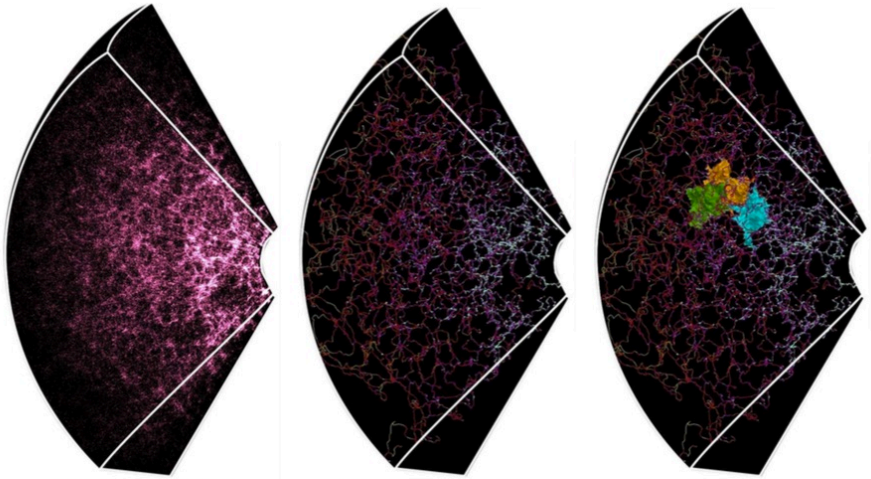
(c) An ascending 3-manifold (*i.e.* a void)



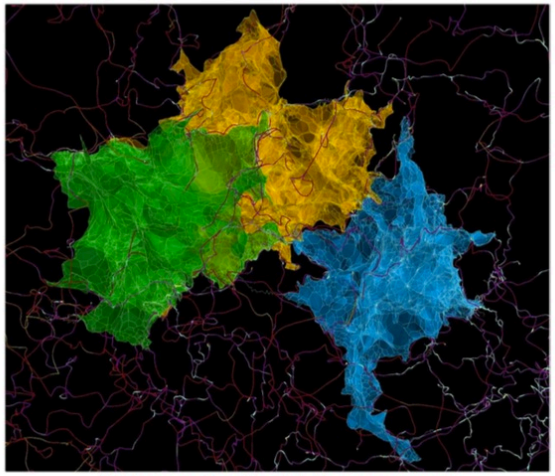
(d) Superposition of an ascending 3-manifold and an ascending 2-manifold on its surface.

Figure 3. An ascending 2-manifold (*i.e.* blue 2D wall) and an ascending 3-manifold (*i.e.* green 3D void) identified in a 512^3 particles $100\, h^{-1}$ Mpc ΛCDM dark matter simulation. The manifolds where computed from a 64^3 particles sub-sample.

Persistent cosmic web (cosmology, Sousbie & al)



(a) A portion of SDSS DR7 (b) The filamentary structure (c) Three voids



(d) A zoom on the voids and the filamentary structure

Figure 10. The detected filamentary structure at a significance level of $5-\sigma$ 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 the the logarithm of the density field. The filaments of the SDSS extracted with DisPerSE is readily available online at the URL `http://www.iap.fr/users/sousbie/`.

[SB1994] S.Barannikov *"The Framed Morse complex and its invariants"* volume 21 of Adv. Soviet Math., pages 93–115. Amer. Math. Soc., Providence, RI, 1994.

[LNV2011] Le Peutrec, D., Nier, F., Viterbo, C. *"Precise Arrhenius Law for p-forms: The Witten Laplacian and Morse–Barannikov Complex"* Ann. Henri Poincaré (2013) 14: 567.

[LSV2018] F. Le Roux, S. Seyfaddini, C. Viterbo, *Barcodes and area-preserving homeomorphisms.* Preprint arXive1810.03139

[Arnold2002] *V.Arnold "What is Mathematics"* MCCME, 2002.– 104 pages. ISBN 978-5-94057-426-2

[ChP] Yu. V. Chekanov, P. E. Pushkar', *"Combinatorics of fronts of Legendrian links and the Arnol'd 4-conjectures" Uspekhi Mat. Nauk, 2005, Vol 60, 1(361), 99–154*

[EH2008] H. Edelsbrunner and J. Harer, *"Persistent homology-a survey."* vol. 453 Cont. Math., 2008, pp. 257–282

[CISZ] G. Carlsson, T. Ishkhanov, V. de Silva, A. Zomorodian,"*On the local behavior of spaces of natural images*" Intern. J. of Comp. Vis., 76 (2008), pp. 1–12

[SPK] *T. Sousbie, C. Pichon, H. Kawahara, "The persistent cosmic web and its filamentary structure II: Illustrations" MNRAS, Vol 414, Issue 1,2011, 384–403,*