Federal State Autonomous Educational Institution
for the Higher Education
National Research University "Higher School of Economics"
Faculty of Mathematics

Zimin Fedor Vladimirovich

Random Complexes and Persistent Homology

Master's thesis

Field of study: 01.04.01 "— Mathematics,

Degree programme: master's educational programme "Mathematics"

Reviewer:
Candidate of Sciences, associate professor
Ilya Ivanovich Ivanov

Scientific supervisor:
Doctor of Sciences, professor
Gorbunov Vasiliy Genadyevich

Moscow 2023

Contents

Introduction

...

1 Background

1.1 Persistent Homology and Giant Cycles

The first we should understand, which objects topological data analysis research.

People call cloud a collection of points $\{x_\alpha\} \subset X$, where X is a metric space. That's interesting to convert that data to some structure, so points can be representated as vertices of some combinatorial graph. And that graph can become scaffold of a simplicial complex. That's a good way to research data, ignoring high dimension of the space [6].

A simplicial complex X (I mean abstract simplicial complex) is a set of vertices $\{v_\alpha\}$ and a collection of its subsets, called simplices (simplex a which is set of elements k have dimension $k-1$ and can be called k-simplex), X such that: for all $a \in X$ for all $b \subset a$ $b \in S$ [5]. The dimension of a simplicial complex is the maximal dimension of it's simplices ($\dim X = \max\limits_{a \in S} \dim a = \max\limits_{a \in S} |a| - 1$).

When simplicial complex is defined, let's continue way to define simplicial homology. Let $\Delta_n(X)$ be the free abelian group with basis on n-simplices e_α^n of X. That groups elements can be rewritten as $\sum_\alpha n_\alpha e_\alpha^n$ and called n-chains. We also can represent them as $\sum_\alpha n_\alpha \sigma_\alpha$ where the $\sigma_\alpha : \Delta^n \to X$ is the characteristic map of .

The boundary of n-simplex $(v_0, ..., v_n)$ is $(n-1)$-simplices $[v_0, ..., v_{i-1}, v_{i+1}, ..., v_n]$. So let the boundary be $\sum_i (-1)^i F_i$. The signs are inserted to take orientations into account, so that all the faces of a simplex are coherently oriented. Using that geometry we can define a boundary homomorphism $\delta : \Delta_n(X) \to \Delta_{n-1}(X)$:

$$\delta_n(\sigma_\alpha) = \sum_i (-1^i)\sigma_\alpha \ : \ [v_0, ..., v_{i-1}, v_{i+1}, ..., v_n]$$

So there is the lemma, which said that the composition $\Delta_n(X) \xrightarrow{\delta_n} \Delta_{n-1}(X) \xrightarrow{\delta_{n-1}}\xrightarrow{\delta_{n-2}}$ $\Delta_{n-2}(X)$ is zero. That's not hard to prove. We have

$$\delta_n(\sigma) = \sum_i (-1)^i \sigma \ : \ [v_0, ..., v_{i-1}, v_{i+1}, ..., v_n]$$

and hence

$$\begin{aligned}\delta_{n-1}\delta_n = &\sum_{j<i} (-1)^i(-1)^j \sigma : [v_0, ..., v_{j-1}, v_{j+1}, ..., v_{i-1}, v_{i+1}, ..., v_n] \\ &+ \sum_{j>i} (-1)^i (-1)^{j-1} \sigma : [v_0, ..., v_{i-1}, v_{i+1}, ..., v_{j-1}, v_{j+1}, ..., v_n] = 0\end{aligned}$$

So we have a sequence of homomorphisms of abelian groups

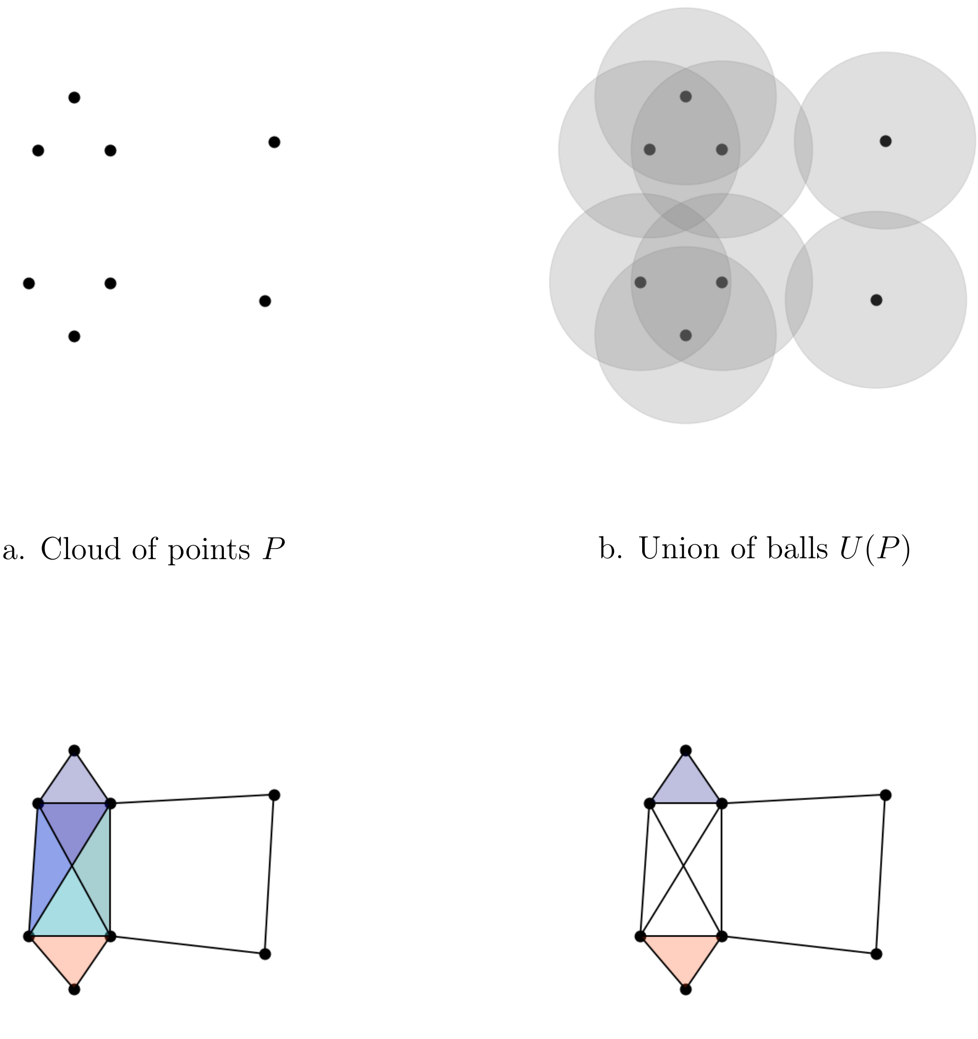
$$\cdots \to C_{n+1} \xrightarrow{\delta_{n+1}} C_n \to \cdots \to C_1 \xrightarrow{\delta_1} C_0 \xrightarrow{\delta_0} 0$$

such that $\delta_n \delta_{n+1} = 0$ for each n. Sequences like that are called chain complexes. Cause the equation $\delta_n\delta_{n+1} = 0$ is equivalent to the inclusion Im $\delta_{n+1} \subset$ Ker δ_n, we can defind the n-th homology group of the chain complex as the quotient group $H_n =$ Ker $\delta_n/$Im δ_{n+1}. In the case of simplicial complex $C_n = \Delta_n(X)$, so the homology group Ker $\delta_n/$Im δ_{n+1} be called the n-th homology group of X and can be noted $H_n^\Delta(X)$. People call the elements of Ker δ_n cycles and the elements of Im δ_{n+1} boundaries [3].

As was said, the basis of H_k corresponds to k-cycles: so the basis of H_0 corresponds to connected components (0-ctvcles), the basis of H_1 to "holes" (1-cycles), the basis of H_2

to "voids" (2-cycles) and more generally H_k represents k-cycles. The dimension of H_k is also known as k-th Betty number ($\beta_k(X) := \dim H_k(X)$), that counts nontrivial k-cycles [11].

One of the natural methods to represent a cloud as a simplicial complex is the Čech complex. [6]. Let $P = \{x_1, x_2, ..., x_n\}$ be a cloud (a collection of points in a metric space (X, ρ)), and let $r \in \mathbb{R}_{>0}$. We wll note the ball of radius r arround x as $B_r(x)$. The Čech complex $C(P, r)$ is constructed such that: The 0-simplices are the points from P; and k-simplex $[x_{i_0}, ..., x_{i_k}]$ is in $C(P, r)$ if the intersection of balls $\cap_{j=0}^{k} B_r(x_{i_j})$ is not empty [8].



a. Cloud of points P

b. Union of balls $U(P)$

c. Čech complex $C(P)$

d. Rips complex $R(P)$

Figure 1: Demonstration of complexes for some cloud

There is associated with Čech complex the union of balls $U(P, r) = \bigcup\limits_{p \in P} B_r(p)$. That's a completely different structures, but there is a Čech theorem (also known as Nerve theorem), which said, that topologically they are very similar [6][1].

Another one natural method to represent a cloud as a simplicial complex is the Rips complex. For a given cloud $P = \{x_\alpha\} \subset (X, \rho)$ the Rips complex R_r is determined by unordered $(k+1)$-tuples of points whose for each pair of points $\{x_\alpha\}_0^k$ the distance between that pairs points less or equal r [6].

Now we can start to talk about persistent homology. That's one of the popuelar tools used in the field of Applied Topological Data Analysis.

Let us X be some topological space. The filtration of X is the set of topological spaces $\{X_t\}_t$ such that $X_s \subset X_t$ for all $s < t$ [11]. Rewriting filtration as $\emptyset = \mathbb{X}_0 \subset \mathbb{X}_1 \subset \cdots \subset \mathbb{X}_m = X$, we can get a linear map for each inclusion

$$0 = H(\mathbb{X}_0) \to H(\mathbb{X}_1) \to \cdots \to H(\mathbb{X}_m) \to H(X)$$

People call the sequences like that persistence modules. Let's split that module into indecomposable sumands like $0 \to F \to \cdots \to F \to 0$, where every nonzero map is the identity. There is a unique such decomposition whose direct sum gives the original module (means which contents all sebsets from a given filtrations). Each summand can be interpreted as the birth of a homology class H^k at its first non-zero term and the death of the same homology class right after its last non-zero term (We will say the birth and the death of k-cycle) [7].

The cycles which never dies (there is no non-zero term after birth for them) we will call giant. But that's possible to define them more formal.

Let us M some topological space and $\{X_t\}$ some filtration such that $X_t \subset M \forall t$. For each t the inclusion map $i: X_t \hookrightarrow M$ induces a map $i_{*,t}: H_k(X_t) \to H_k(M)$. The image of that map $i_{*,t}$ stands for all the cycles that exists in X_t and are mapped to nontrivial cycles in M. These are giant cycles [11].

Let's throw few examples of persistent homology: we have discussed Cech complex (C_r), Reese complex (R_r) and associated with the Cech the union of balls U_r. So using them we can defind filtrations $\{U_r(P)\}_{r \in \mathbb{R}_+}$, $\{C_r(P)\}_{r \in \mathbb{R}_+}$, $\{R_r(P)\}_{r \in \mathbb{R}_+}$ for a given cloud $P \subset X$, where X is some metric space. You can see example in Figure 2.

...

1.2 Lattices, Voronoi Cells and their Extropolation on Torus

In this chapter we will show default definitions about lattices in $\mathbb{R}^n$, talk about Voronoi cells and then extropolate their definitions to the torus case.

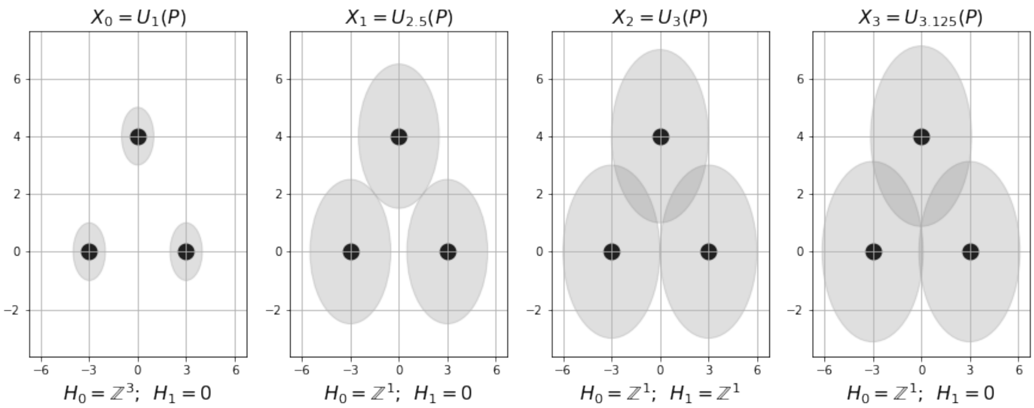
A lattice in $\mathbb{R}^n$ is a subset $\Gamma \subset \mathbb{R}^n$ with the property that there exists a basis $(e_1, ..., e_n)$ of $\mathbb{R}^n$ s.t. $\Gamma = \mathbb{Z}e_1 \oplus \cdots \oplus \mathbb{Z}e_n$. The fundamental parallelotope of lattice Γ is $\{\lambda_1 e_1 + \cdots + \lambda_n e_n \mid 0 \le \lambda_i \le 1\}$.

The matrix with lines $e_1, ..., e_n$ is called generator matrix of lattice Γ.The matrix of scalar products

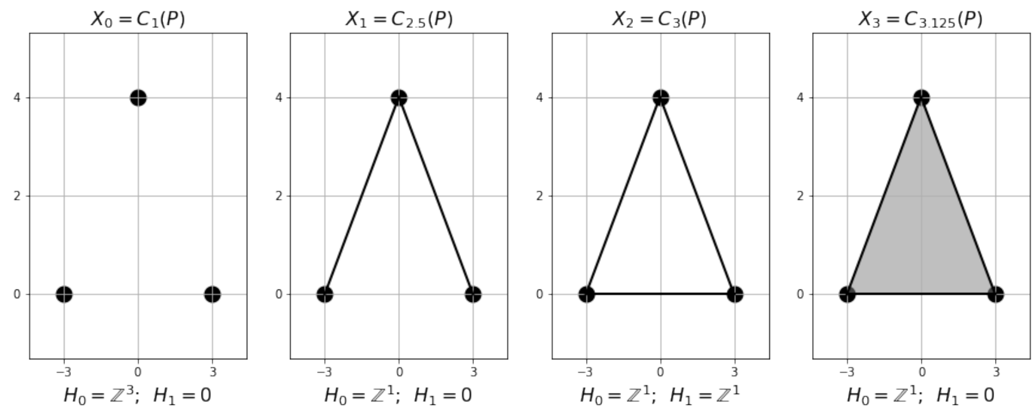
$$G = \begin{pmatrix} \langle e_1, e_1 \rangle & \langle e_1, e_2 \rangle & \langle e_1, e_n \rangle & \\ \langle e_2, e_1 \rangle & \langle e_2, e_2 \rangle & \langle e_2, e_n \rangle & \\ \cdots & \cdots & \cdots & \cdots \\ \langle e_n, e_1 \rangle & \langle e_n, e_2 \rangle & \langle e_n, e_n \rangle & \end{pmatrix}$$

is known as Gram matrix [4].

If one lattice can be obtained from another by (possibly) a rotation, reflection and change of scale we say they are equivalent. More formally two lattices with generator matrices G and G' are equivalent, if and only if $G' = cUMB$, where c is some nonzero



a. Filtration $\{U_r(P)\}_{r>0}$



b. Filtration $\{C_r(P)\}_{r>0}$

Figure 2: There are filtrations $\{U_r(P)\}_{r>0}, \{C_r(P)\}_{r>0} \subset \mathbb{R}^2$ where $P = \{(-3,0),\ (3,0),\ (0,4)\}$. In pictures a and b tou can see that filtrations for parameters $r \in \{1,\ 2.5,\ 3,\ 3.125\}$. All them except the first are birth or death values.

constant, U is a matrix with integer entries s.t $|\det U| = 1$ and B is a real orthogonal matrix. The corresponding Gram matrices are relited by $A' = c^2 U A U^T$ [2].

The binary operation for two vectors α, β $S_\alpha(\beta)$ is called reflection if $S_\alpha(\alpha) = -\alpha$ and $S_\alpha(\beta) = \beta$ for each $\beta \perp \alpha$. Usually $S_\alpha(\beta) = \beta - 2\dfrac{\langle \lambda, \beta \rangle}{\langle \lambda, \alpha \rangle}\alpha$.

The finite set of vetcors Φ is a root system if $\Phi \cap \mathbb{R}\alpha = \{\alpha, -\alpha\}$ and $S_\alpha \Phi = \Phi$ for each α from Φ [9].

The lattices generated by root systems called root lattices.

Let's throw few examples of lattices, which will be interesting in this work. That's root lattices.

The Lattice $\mathbb{Z}^n$ is the most simpliest lattice, that's just the set of all vectors from $\mathbb{R}^n$ which all elements are integer.

$$\mathbb{Z}^n = \{(x_1, ..., x_n) :\ x_i \in \mathbb{Z}\} \subset \mathbb{R}^n$$

That's generator matrix and Gram matrix is identity matrix $\mathbb{I}_n$.

The Lattice A_n $(n \geq 1)$ is the subgroup of lattice $\mathbb{Z}^{n+1}$ which elements lie on the hyperplane $\sum\limits_{i=0}^{n} x_i = 0$.

$$A_n = \{(x_1, x_1, ..., x_n) \in \mathbb{Z}^n : \ x_1 + \cdots x_n = 0\}$$

That's generator matrix can look like that:

$$M_{A_n} = \begin{pmatrix} -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

And there are two types of gram matrices possible:

$$\begin{pmatrix} 2 & -1 & 0 & \cdots & 0 & 0 \\ -1 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}$$

or

$$\begin{pmatrix} 2 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 1 & \cdots & 1 & 1 \\ 1 & 1 & 2 & \cdots & 1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 1 & 1 & \cdots & 2 & 1 \\ 1 & 1 & 1 & \cdots & 1 & 2 \end{pmatrix}$$

The Lattice D_n can be defined for $n \geq 3$. That's the subgroup of $\mathbb{Z}^n$ consists all elements which sum coefficients is even.

$$D_n = \{(x_1, ..., x_n) \in \mathbb{Z}^n : \ x_1 + \cdots + x_n \mid 2\}$$

That's generator matrix can look like that :

$$M_{D_n} = \begin{pmatrix} -1 & -1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & -1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & -1 & 1 \end{pmatrix}$$

and that's Gram matrix:

$$\begin{pmatrix} 2 & 2 & 0 & \cdots & 0 & 0 \\ 2 & 2 & -1 & \cdots & 0 & 0 \\ 0 & -1 & 2 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 2 & -1 \\ 0 & 0 & 0 & \cdots & -1 & 2 \end{pmatrix}$$

People also call that lattice checkerboard lattice [2].

That's easy to see, that's $A_3 \equiv D_3$, cause by permutations of vectors in M_{A_3} and calculating corresponding Gram matrix we will get the Gram matrix of D_3:

$$\begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

Let's define a lattice Γ^* dual to Γ as $\{x \in \mathbb{R}^{\ltimes} : \ x \cdot y \in \mathbb{Z} \, \forall y \in \Gamma\}$ [4].

That's easy to understand, that lattice $\mathbb{Z}^n$ is self-dual (self-dual lattice is the lattice Γ such that $\Gamma^* = \Gamma$)

The lattice A_n^* dual to A_n has folows generator matrix

$$M_{A_n^*} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 1 & 0 & -1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 1 & 0 & 0 & \cdots & -1 & 0 \\ \frac{-n}{n+1} & \frac{1}{n+1} & \frac{1}{n+1} & \cdots & \frac{1}{n+1} & \frac{1}{n+1} \end{pmatrix}$$

And the related to that definition Gram matrix will be

$$\begin{pmatrix} n & -1 & -1 & \cdots & -1 \\ -1 & n & -1 & \cdots & -1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -1 & -1 & -1 & \cdots & n \end{pmatrix}$$

That's not hard to check that $M_{A_n} \cdot M_{A_n^*}^T = \mathbb{I}_n$.

The generator matrix of lattice D_n^* dual to D_n will be

$$M_{D_n^*} = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & \cdots & \frac{1}{2} & \frac{1}{2} \end{pmatrix}$$

D_3^* is the body centered cubic lattice, like A_3^*, and $D_4^* \equiv D_4$ [2].

...

Let's define d-dimensional torus as $\mathbb{R}^n/\mathbb{Z}^n$ or $(\mathbb{R}/\mathbb{Z})^n$ (Omer Bobrowski and Shmueluse Weinberger use that definition in [8]). Not hard to see, that $\mathbb{R}/v_1\mathbb{Z} \times \cdots \times \mathbb{R}/v_n\mathbb{Z}$ $(v_1, ..., v_n \in \mathbb{R}_{>0})$ will be the homeomorphically-same object.

Let's redefine lattice thinking, they lie not just on $\mathbb{R}^n$, but on some torus with defined equivalence relation given by vector $v = (v_1, ..., v_n) \in \mathbb{R}^n_{>0}$: for $a, b \in \mathbb{R}$ $a \equiv b$ if $a - b = vk$ $k \in \mathbb{Z}$. That's litterally equivalence relation from the given torus definition. If for the lattice basis $e_1, ..., e_n$ and the given vector v and eacn $k \in \{1, ..., n\}$ there exists $j \in \mathbb{Z}$ s.t. $jv_k \mid e_{ik}$ the lattice will have finite number of elemens on torus defined by vector v.

...

2 ...

2.1 Percoation on Cells

Continuum-percolation models are random processes in which subsets $D_i \subseteq X$ where $i \in \{1, ..., N\}$ are chosen randomly with some probability distribution in a structure $\bigcup_{i=1}^{N} D_i$ [10].

...

References

[1] Karol Borsuk. (1948) On the imbedding of systems of compacta in simplicial complexes. Fundamenta Mathematicae, 35(1):217–234, 1948.

[2] Conway, John Sloane, N.. (1988). Sphere Packings, Lattices and Groups. 10.1007/978-1-4757-2016-7.

[3] Hatcher, A. (2001). Algebraic topology. Proceedings of The Edinburgh Mathematical Society - PROC EDINBURGH MATH SOC. 46. 511-512. 10.1017/S0013091503214620.

[4] Ebeling, Wolfgang. (2002). Lattices and Codes. 10.1007/978-3-322-90014-2.

[5] Prasolov, V. V. (2006), Elements of combinatorial and differential topology, American Mathematical Society, ISBN 0-8218-3809-1, MR 2233951

[6] Ghrist, Robert. (2008). Barcodes: The persistent topology of data. BULLETIN (New Series) OF THE AMERICAN MATHEMATICAL SOCIETY. 45. 10.1090/S0273-0979-07-01191-3.

[7] Edelsbrunner, Herbert Morozov, D.. (2012). Persistent homology: Theory and practice. Proceedings of the European Congress of Mathematics. 31-50.

[8] Bobrowski, Omer Weinberger, Shmuel. (2015). On the Vanishing of Homology in Random Čech Complexes. Random Structures and Algorithms. 10.1002/rsa.20697.

[9] Е.Ю. Смирнов "Группы отражений и правильные многогранники" Издание второе, исправленное и дополненное. Москва. МЦНМО. 2018.

[10] Speidel, Leo Harrington, Heather Chapman, Stephen Porter, Mason. (2018). Topological data analysis of continuum percolation with disks. Physical Review E. 98. 10.1103/PhysRevE.98.012318.

[11] Bobrowski, Omer Skraba, Primoz. (2019). Homological Percolation and the Euler Characteristic.