Arnold Math J. (2017) 3:197–204
DOI 10.1007/s40598-016-0062-6

RESEARCH CONTRIBUTION

Random Chain Complexes

Viktor L. Ginzburg[1] · **Dmitrii V. Pasechnik**[2]

Received: 16 March 2016 / Revised: 9 December 2016 / Accepted: 23 December 2016 /
Published online: 25 January 2017

Abstract We study random, finite-dimensional, ungraded chain complexes over a finite field and show that for a uniformly distributed differential a complex has the smallest possible homology with the highest probability: either zero or one-dimensional homology depending on the parity of the dimension of the complex. We prove that as the order of the field goes to infinity the probability distribution concentrates in the smallest possible dimension of the homology. On the other hand, the limit probability distribution, as the dimension of the complex goes to infinity, is a super-exponentially decreasing, but strictly positive, function of the dimension of the homology.

Keywords Random chain complexes · Homology · Floer theory

Mathematics Subject Classification 05E99 · 55U15 · 53D99 · 60D99

The work is partially supported by NSF grant DMS-1308501 (VG) and the EU Horizon 2020 research and innovation programme, grant agreement OpenDreamKit No. 676541 (DP).

✉ Viktor L. Ginzburg
ginzburg@ucsc.edu

Dmitrii V. Pasechnik
dmitrii.pasechnik@cs.ox.ac.uk

1 Department of Mathematics, UC Santa Cruz, Santa Cruz, CA 95064, USA

2 Department of Computer Science, University of Oxford, Wolfson Building, Parks Road, Oxford OX1 3QD, UK

1 Introduction

We study random, finite-dimensional, ungraded chain complexes over a finite field and we are interested in the probability that such a complex has homology of a given dimension. We show that for a uniformly distributed differential the complex has the smallest possible homology with the highest probability.

To be more specific, consider an n-dimensional vector space V over a finite field $\mathbb{F} = \mathbb{F}_q$ of order q and let D be a differential on V, i.e., a linear operator $D\colon V \to V$ with $D^2 = 0$. We are interested in the probability $p_r(q, n)$ with which a chain complex (V, D) has homology $\ker D/\operatorname{im} D$ of a given dimension r for fixed n and q. The differential D is uniformly distributed and $p_r(q, n)$ is simply the ratio $c_r(q, n)/c(q, n)$, where $c_r(q, n)$ is the number of complexes with r-dimensional homology and $c(q, n)$ is the number of all complexes. The numbers $c_r(q, n)$ are explicitly calculated by Kovacs (1987); see also Theorem 2.2.

We mainly focus on large complexes, i.e., on the limits as q or n go to infinity. Clearly, r and n must have the same parity and we separately analyze the asymptotic behavior of the sequence $p_0(q, n), p_2(q, n), \ldots$, where n is even, and the sequence $p_1(q, n), p_3(q, n), \ldots$ for n odd.

As $q \to \infty$ with n fixed, the probability concentrates in the lowest possible dimension, i.e., $p_0(q, n) \to 1$ or $p_1(q, n) \to 1$ depending on the parity of n, while $p_r(q, n) \to 0$ for $r > 1$. This is consistent with the observation that over $\mathbb{C}$ and even $\mathbb{R}$ (see Lemma 3.1) a generic complex has 0- or 1-dimensional homology, i.e., that such complexes form the highest dimensional stratum in the variety of all n-dimensional complexes. Indeed, one can expect the probability distributions for large q to approximate the generic situation in zero characteristic. We do not know, however, if the density functions converge in any sense as $q \to \infty$ to some probability density on the variety of n-dimensional complexes over, e.g., $\mathbb{R}$.

When q and r are fixed and $n \to \infty$ through either even or odd integers depending on the parity of r, the situation is more subtle. In this case, all limit probabilities $p_r(q) = \lim_{n\to\infty} p_r(q, n)$ are positive. However, the sequences $p_0, p_2, \ldots$ and $p_1, p_3, \ldots$ are super-exponentially decreasing and for a large q all terms in these sequences but the first one are very close to zero while the first is then, of course, close to 1. When $q = 2$ and r is even, we have $p_0 \approx 0.6$, $p_2 \approx 0.4$, $p_4 \approx 0.0075$ and other terms are very small. We explicitly calculate the ratios $p_r(q)/p_0(q)$ and $p_r(q)/p_1(q)$ and p_0 and p_1 in Theorem 2.1.

The proofs of these facts are elementary and quite simple. However, we have not been able to find in the literature any probability calculations in this basic case where random chain complexes are stripped of all additional structures including a grading. (The combinatorial part of our proof, Theorem 2.2, is contained in [Kovacs (1987), Lemma 5].) In contrast, random complexes of geometrical origin and underlying random geometrical and topological objects have been studied extensively and from various perspectives. Among such random objects are, for instance, random simplicial complexes of various types [see (Aronshtam et al. 2013; Bobrowski and Kahle 2014; Costa and Farber 2014, 2015; Kahle 2011; Meshulam 2013; Meshulam and Wallach 2009; Pippenger and Schleich 2006; Yogeshwaran et al. 2014) and references therein] and random Morse functions [see, e.g., (Arnold 2006, 2007; Collier et al. 2017; Nicolaescu 2012)].

These works utilize several models of randomness all of which appear to be quite different from the one, admittedly rather naive, used here. This makes a direct comparison difficult. One way to interpret our result is that, for a large complex, sufficiently non-trivial homology is indicative of some structure, a constraint limiting randomness. Note that such a structure can be as simple as a $\mathbb{Z}$-grading confined to a fixed range of degrees. A dimensional constraint of this type is usually inherent in geometrical complexes, and it would be interesting to analyze its effect (if any) on the probability distribution in our purely algebraic setting. Another consequence of the result is that the assertion that a complex has large homology carries more information than the assertion that it has small homology.

The main motivation for our setting comes from Hamiltonian Floer theory for closed symplectic manifolds; see, e.g., Salamon (1999) and references therein. A Hamiltonian diffeomorphism is the time-one map of the isotopy generated by a time-dependent Hamiltonian. To such a diffeomorphism one can associate a certain complex, called the Floer complex, generated by its fixed points or, equivalently, the one-periodic orbits of the isotopy. Hence the dimension of the Floer homology gives a lower bound for the number of one-periodic orbits. The homology is independent of the Hamiltonian diffeomorphism. In addition, one can fix the free homotopy class of the orbits. (This construction is similar to Morse theory and, in fact, Floer theory is a version of Morse theory for the action functional.)

In many instances, e.g., often generically or for all symplectic manifolds with vanishing first Chern class such as tori, the dimension of the Floer complex grows with the order of iteration of the diffeomorphism; see Ginzburg and Gürel (2015). In other words, the complex gets larger and larger as time in this discrete dynamical system grows. Moreover, the differential in the complex is usually impossible to describe explicitly, and hence it makes sense to compare the behavior of the complex and its homology with the generic or random situation. The Floer homology for contractible periodic orbits is isomorphic to the homology of the underlying manifold. Therefore, by our result, even though the Floer complex appears to be very "noisy" for large iterations and random on a bounded action scale, it has large homology groups and is actually very far from random. For non-contractible orbits, the dimension of the Floer complex is also known to grow in many settings; see Ginzburg and Gürel (2015); Gürel (2013). However, in this case the Floer homology is zero and the complex may well be close to random. Note also that in some instances the Floer complex is $\mathbb{Z}$-graded, but the grading is not supported within any specific interval of degrees. Moreover, in contrast with geometrical random complexes, the grading range of the Floer homology usually grows with the order of iteration (Salamon and Zehnder 1992), and while it is not clear how to correctly account for an unbounded grading in a random model, such a grading is unlikely to affect the probability distribution.

One aspect of Floer theory which is completely ignored in our model is the action filtration. This filtration is extremely important and, in particular, it allows one to treat Floer theory in the context of persistent homology and topological data analysis; see Carlsson (2009); Ghrist (2008). This connection has recently been explored in

Polterovich and Shelukhin (2014); Usher and Zhang (2015). However, it is not entirely clear how to meaningfully incorporate the action filtration into our model.

2 Main Results

Let, as in the introduction, (V, D) be an ungraded n-dimensional chain complex with differential D over a finite field $\mathbb{F} = \mathbb{F}_q$ of order q. In other words, $V = \mathbb{F}^n$ and D is a linear operator on V with $D^2 = 0$. We denote by $c(q, n)$ the number of such complexes, i.e., the number of differentials D. The dimension r of the homology $\ker D/\operatorname{im} D$ has the same parity as n and we let $c_r(q, n)$ be the number of complexes with homology of dimension r. (In what follows, we always assume that r and n have the same parity.) Clearly,

$$c(q, n) = c_0(q, n) + c_2(q, n) + \cdots + c_n(q, n)$$

when n is even and

$$c(q, n) = c_1(q, n) + c_3(q, n) + \cdots + c_n(q, n)$$

when n is odd.

Furthermore, denote by

$$p_r(q, n) = \frac{c_r(q, n)}{c(q, n)}$$

the probability (with respect to the uniform distribution) of a complex to have r-dimensional homology. Our main result describes the behavior of $p_r(q, n)$ as the size of the complex, i.e., q or n, goes to infinity.

Theorem 2.1 *Let $p_r(q, n)$ be as above.*

(i) *For a fixed n, we have*

$$\lim_{q\to\infty} p_r(q, n) = 0 \text{ when } r > 1,$$

and $p_0(q, n) \to 1$ when n is even and $p_1(q, n) \to 1$ when n is odd as $q \to \infty$.

(ii) *For a fixed q and r, the limits*

$$p_r(q) = \lim_{n\to\infty} p_r(q, n)$$

exist and $0 < p_r(q) < 1$ for all q and r. Furthermore, when $r \geq 2$ is even, we have

$$\frac{p_r(q)}{p_0(q)} = \frac{q^{r/2}}{\prod_{j=1}^{r}(q^j - 1)} \tag{2.1}$$

and

$$p_0(q) = \frac{1}{1+S}, \quad \textit{where} \quad S = \sum_{k=1}^{\infty} \frac{q^k}{\prod_{j=1}^{2k}(q^j - 1)}. \tag{2.2}$$

When $r \geq 3$ is odd,

$$\frac{p_r(q)}{p_1(q)} = \frac{(q-1)q^{(r-1)/2}}{\prod_{j=1}^{r}(q^j - 1)}$$

and

$$p_1(q) = \frac{1}{1+S'}, \quad \textit{where} \quad S' = (q-1)\sum_{k=1}^{\infty} \frac{q^k}{\prod_{j=1}^{2k+1}(q^j - 1)}.$$

The proof of this theorem is based on an explicit calculation of $c_r(q, n)$. To state the result, denote by $\mathrm{GL}_k(q)$ the general linear group of $k \times k$ invertible matrices over $\mathbb{F}_q$ and recall that

$$|\,\mathrm{GL}_k(q)| = q^{k(k-1)/2} \prod_{j=1}^{k}(q^j - 1).$$

Then we have the following particular case of [Kovacs (1987), Lemma 5].

Theorem 2.2 (Kovacs 1987) *Let as above $c_r(q, n)$ be the number of n-dimensional complexes over $\mathbb{F}_q$ with homology of dimension r. Then*

$$c_r(q, n) = \frac{|\,\mathrm{GL}_n(q)|}{|\,\mathrm{GL}_m(q)| \cdot |\,\mathrm{GL}_r(q)| \cdot q^{2mr+m^2}}, \tag{2.3}$$

where $2m + r = n$.

Even though this result is not new, for the sake of completeness we include its proof, which is very simple and short, in the next section.

Remark 2.3 We do not have simple expressions for the probabilities $p_r(q, n)$ and the total number of complexes $c(q, n)$. However, when $q = 2$, the differentials D are in one-to-one correspondence with involutions of $\mathbb{F}_2^n$. (An involution necessarily has the form $I + D$ and, as is easy to see, different differentials D give rise to different involutions.) Hence, $c(2, n)$ is equal to the number of involutions. This number is expressed in Fulman and Vinroot (2014) via a generating function and an asymptotic formula for $c(q, n)$ has been recently obtained in Fulman et al. (2016). It is possible that at least when $q = 2$ our probability formulas can be further simplified using the results from those papers.

3 Proofs

The proof of Theorem 2.2 is based on the observation that the differential in a finite-dimensional complex over any field $\mathbb{F}$ can be brought to its Jordan normal form or, equivalently, a complex over $\mathbb{F}$ can be decomposed into a sum of elementary complexes, i.e., into a sum of two-dimensional complexes with zero homology and one-dimensional complexes. To be more precise, we have the following elementary observation.

Lemma 3.1 *Let V be a finite-dimensional vector space over an arbitrary field $\mathbb{F}$ and let $D\colon V \to V$ be an operator with $D^2 = 0$. Then, in some basis, D can be written as a direct sum of 1×1 and 2×2 Jordan blocks with zero eigenvalues.*

When $\mathbb{F}$ is algebraically closed, this follows immediately from the Jordan normal form theorem. Hence, the emphasis here is on the fact that the field $\mathbb{F}$ is immaterial. For the sake of completeness, we outline a proof of the lemma.

Proof Let us pick an arbitrary basis $\{e_1, \dots, e_m\}$ of $\operatorname{im} D$ and extend it to a basis of $\ker D \supset \operatorname{im} D$ by adding elements $\{f_1, \dots, f_r\}$. Furthermore, pick arbitrary vectors e'_i with $De'_i = e_i$. Then $\{e'_1, e_1, \dots, e'_m, e_m, f_1, \dots, f_r\}$ is the required basis of V. □

Proof of Theorem 2.2 Let D be a differential on an n-dimensional vector space V over a finite field $\mathbb{F} = \mathbb{F}_q$. Assume that the homology of the complex (V, D) is r-dimensional. By Lemma 3.1, D is conjugate to the map D_r which is the direct sum of r 1×1 zero blocks and m 2×2 Jordan blocks with zero eigenvalues, where $2m + r = n$.

Let C_r be the centralizer of D_r in $\mathrm{GL}_n(q)$. The complexes with r-dimensional homology are in one-to-one correspondence with $\mathrm{GL}_n(q)/C_r$. Thus, to prove (2.3), it suffices to show that

$$|C_r| = |\,\mathrm{GL}_m(q)| \cdot |\,\mathrm{GL}_r(q)| \cdot q^{2mr+m^2}. \tag{3.1}$$

The elements of C_r are $n \times n$ invertible matrices $X \in \mathrm{GL}_n(q)$ commuting with D_r. In what follows, it is convenient to work with the basis $e_1, \dots, e_m, f_1, \dots, f_r, e'_1, \dots, e'_m$ in the notation from the proof of Lemma 3.1. Thus we can think of X as a 3×3-block matrix with $m \times m$ block X_{11}, the block X_{12} having size $m \times r$, and X_{13} being again $m \times m$, etc. In the same format, D_r is then the matrix with only one non-zero block. This is the top-right corner $m \times m$-block, which is I. Then, as a straightforward calculation shows, the commutation relation $XD_r = D_r X$ amounts to the conditions that $X_{11} = X_{33}$, and $X_{21} = 0$, $X_{31} = 0$ and $X_{32} = 0$. In particular, X is an upper block-triangular matrix. Hence, X is invertible if and only if $X_{11} = X_{33}$ and X_{22} are invertible. There are no constraints on the remaining blocks X_{12}, X_{13} and X_{23}. Now (3.1) follows. □

Proof of Theorem 2.1 Throughout the proof, we assume that r and n are even. The case where these parameters are odd can be handled in a similar fashion.

As the first step, we express $p_r(q, n)/p_0(q, n)$ explicitly. Clearly,

$$\frac{p_r(q, n)}{p_0(q, n)} = \frac{c_r(q, n)}{c_0(q, n)} = \frac{|C_0|}{|C_r|}.$$

Using (2.3) or (3.1) and tidying up the resulting expression, we have

$$\begin{aligned}\frac{p_r(q,n)}{p_0(q,n)} &= \frac{q^{n^2/4}\cdot q^{n(n/2-1)/4}\cdot \prod_{j=1}^{n/2}(q^j-1)}{q^{2mr+m^2}\cdot q^{m(m-1)/2}\cdot q^{r(r-1)/2}\cdot \prod_{j=1}^{r}(q^j-1)\cdot \prod_{j=1}^{m}(q^j-1)}\\ &= \frac{\prod_{j=m+1}^{m+r/2}(q^j-1)}{q^{mr/2+r(r/2-1)/4}\cdot \prod_{j=1}^{r}(q^j-1)},\end{aligned}$$

where as above $n = 2m + r$ and $r \geq 2$, which we can then rewrite as

$$\frac{p_r(q,n)}{p_0(q,n)} = \frac{q^{r/2}}{\prod_{j=1}^{r}(q^j-1)}\cdot \prod_{j=1}^{r/2}\left(1-\frac{1}{q^{m+j}}\right). \tag{3.2}$$

Now it is clear that

$$\frac{p_r(q,n)}{p_0(q,n)} \sim q^{-r^2/2} \quad \text{as} \quad q\to\infty$$

with $r \geq 2$ and n fixed. In particular, this ratio goes to zero as $q \to \infty$. The number of the terms in the sum

$$\sum_j p_j(q,n) = 1$$

with j ranging through even integers from 0 to n is equal to $n/2 + 1$ and thus this number is independent of q. Hence, $p_0(q,n) \to 1$ and $p_r(q,n) \to 0$ when $r \geq 2$ as $q \to \infty$. This proves the first assertion of the theorem.

To prove the second part, first note that by (3.2)

$$\frac{p_r(q,n)}{p_0(q,n)} \to \frac{q^{r/2}}{\prod_{j=1}^{r}(q^j-1)} \tag{3.3}$$

as $m \to \infty$ or, equivalently, $n \to \infty$ with r and q fixed.

Furthermore, in a similar vein, it is not hard to show that

$$\sum_{r>0}\frac{p_r(q,n)}{p_0(q,n)} \to S := \sum_r \frac{q^{r/2}}{\prod_{j=1}^{r}(q^j-1)} \quad \text{as } n\to\infty,$$

where, on the left, the sum is taken over all even integers from 2 to n and, on the right, the sum is over all even integers $r \geq 2$. Therefore, letting $n \to \infty$ in the identity

$$1+\sum_{r>0}\frac{p_r(q,n)}{p_0(q,n)} = \frac{1}{p_0(q,n)},$$

we conclude that the limit $p_0(q) = \lim_{n\to\infty} p_0(q, n)$ exists and $p_0(q) = 1/(1 + S)$, which proves (2.2). Now, by (3.3), the limits $p_r(q) = \lim_{n\to\infty} p_r(q, n)$ for $r \geq 2$ also exist, and hence (2.1) holds. This completes the proof of the theorem. □

Remark 3.2 The sequence $p_r(q, n)$ is decreasing as a function of r. This readily follows from (3.2).

Acknowledgements The authors are grateful to Robert Ghrist, Başak Gürel, Jiang-Hua Lu, Roy Meshulam and Leonid Polterovich for useful discussions and comments. The authors would also like to thank the referee for pointing out [Kovacs (1987), Lemma 5] to them. A part of this work was carried out while the second author was visiting the Simons Institute for the Theory of Computing and he would like to thank the institute for its warm hospitality.

References

Arnold, V.I.: Smooth functions statistics. Funct. Anal. Other Math. **1**, 111–118 (2006)
Arnold, V.I.: Topological classification of Morse functions and generalisations of Hilbert's 16-th problem. Math. Phys. Anal. Geom. **10**, 227–236 (2007)
Aronshtam, L., Linial, N., Łuczak, T., Meshulam, R.: Collapsibility and vanishing of top homology in random simplicial complexes. Discrete Comput. Geom. **49**, 317–334 (2013)
Bobrowski, O., Kahle, M.: Topology of random geometric complexes: a survey (2014, Preprint). arXiv:1409.4734
Carlsson, G.: Topology and data. Bull. Am. Math. Soc. (N.S.) **46**, 255–308 (2009)
Collier, B., Hockensmith, D.L., Kerman, E., Wong, Y.-S.: Realistic Morse complexes (2017, in preparation)
Costa, A., Farber, M.: Random simplicial complexes (2014, Preprint). arXiv:1412.5805
Costa, A.E., Farber, M.: Geometry and topology of random 2-complexes. Isr. J. Math. **209**, 883–927 (2015)
Fulman, J., Guralnick, R., Stanton, D.: Asymptotics of the number of involutions in finite classical groups (2016, Preprint). arXiv:1602.03611
Fulman, J., Vinroot, C.R.: Generating functions for real character degree sums of finite general linear and unitary groups. J. Algebraic Comb. **40**, 387–416 (2014)
Ghrist, R.: Barcodes: the persistent topology of data. Bull. Am. Math. Soc. (N.S.) **45**, 61–75 (2008)
Ginzburg, V.L., Gürel, B.Z.: Non-contractible periodic orbits in Hamiltonian dynamics on closed symplectic manifolds. Compos. Math. (2015, Preprint). arXiv:1503.07145
Ginzburg, V.L., Gürel, B.Z.: The Conley conjecture and beyond. Arnold Math. J. **1**, 299–337 (2015)
Gürel, B.Z.: On non-contractible periodic orbits of Hamiltonian diffeomorphisms. Bull. Lond. Math. Soc. **45**, 1227–1234 (2013)
Kahle, M.: Random geometric complexes. Discrete Comput. Geom. **45**, 553–573 (2011)
Kovacs, A.: Some enumeration problems for matrices over a finite field. Linear Algebra Appl. **94**, 223–236 (1987)
Meshulam, R.: Bounded quotients of the fundamental group of a random 2-complex (2013, Preprint). arXiv:1308.3769
Meshulam, R., Wallach, N.: Homological connectivity of random k-dimensional complexes. Random Struct. Algorithms **34**, 408–417 (2009)
Nicolaescu, L.I.: Random Morse functions and spectral geometry (2012, Preprint). arXiv:1209.0639
Pippenger, N., Schleich, K.: Topological characteristics of random triangulated surfaces. Random Struct. Algorithms **28**, 247–288 (2006)
Polterovich, L., Shelukhin, E.: Autonomous Hamiltonian flows, Hofer's geometry and persistence modules (2014, Preprint). arXiv:1412.8277
Salamon, D.A.: Lectures on Floer homology. In: Symplectic Geometry and Topology, IAS/Park City Math. Ser., vol. 7, pp. 143–229. Amer. Math. Soc., Providence (1999)
Salamon, D., Zehnder, E.: Morse theory for periodic solutions of Hamiltonian systems and the Maslov index. Commun. Pure Appl. Math. **45**, 1303–1360 (1992)
Usher, M., Zhang, J.: Persistent homology and Floer–Novikov theory (2015, Preprint). arXiv:1502.07928
Yogeshwaran, D., Subag, E., Adler, R.J.: Random geometric complexes in the thermodynamic regime (2014, Preprint). arXiv:1403.1164