

ON COHOMOLOGY THEORY OF (DI)GRAPHS

AN HUANG AND SHING-TUNG YAU

(communicated by Donald M. Davis)

Abstract

To a digraph with a choice of certain integral basis, we construct a CW complex, whose integral singular cohomology is canonically isomorphic to the path cohomology of the digraph that was studied recently. The homotopy type of the CW complex turns out to be independent of the choice of basis. The construction is functorial, and it makes many of the recently proved properties of digraph cohomology and homotopy manifest. Furthermore, one gets an expected formula for the cup product of forms on a digraph. On the other hand, we present an approach using sheaf theory to reformulate (di)graph cohomologies. The investigation of the digraph path cohomology from this sheaf theory framework leads to a subtle version of Poincaré lemma for digraphs, which follows from the construction of the CW complex.

1. Introduction

In the past few years, there has been rapidly growing interest in developing geometric concepts in the context of graphs beyond spectral graph theory. See, for example, [6] for a short exposition. In particular, there exist several attempts to define the homology and cohomology of (di)graphs—for example, via cliques [2] or via path algebra [9, 3]. One potential application of such ideas concerns topological phases in condensed matter physics; see, for example, [11].

Our first purpose here is to try to better understand the path cohomology of digraphs [9]. This is an interesting theory that is expected to play the role of singular cohomology, or in some nice cases de Rham cohomology, for digraphs. Chapters 5, 6, and 7 of [9] discuss some of its nice but perhaps combinatorially subtle properties, and more basic properties regarding homotopy are proved in [10], by applying ideas of traditional algebraic topology to digraphs. In this paper, we construct a CW complex from a digraph with a given choice of integral basis of a certain kind, whose integral singular cohomology is canonically isomorphic to the path cohomology of digraphs. We will see that this construction gives rise to a functor from the category of digraphs to a skeleton of the homotopy category of CW complexes, preserving

Received April 14, 2015, revised June 9, 2015; published on December 3, 2015.

2010 Mathematics Subject Classification: 05C10, 55P65.

Key words and phrases: digraph cohomology, CW complex, homotopy, sheaf cohomology for (di)graphs.

Article available at <http://dx.doi.org/10.4310/HHA.2015.v17.n2.a18>

Copyright © 2015, International Press. Permission to copy for private use granted.

products. Some immediate consequences of the existence of this functor include that the path cohomology of digraphs is homotopy invariant, that the Künneth formula holds, and that there exists a functorial cup product on the path cohomology that can be lifted to the level of forms, as [10, 9] proved. Furthermore, one can then define arbitrary higher homotopy groups for a digraph to be those of the CW complex, as in [10], but in a possibly different way. In addition, we get a very simple formula for the cup product of forms on the digraph, which is actually implicitly contained in [10], but here we understand the formula in a more geometric way. We hope these results will be the beginning of a systematic investigation, exploring the connection between digraph theory and topology. This construction may be thought of as a generalization of associating a simplicial complex to a graph, but is much more subtle, and has better categorical behavior. Intuitively, it may be viewed as, in some sense, an inverse construction to a particular generalized concept of triangulation of a manifold, which we hope to investigate in future works.

As the combinatorial Laplacian is a central object in (di)graph theory, one clear motivation for developing (di)graph cohomology theories is, in particular, to get interesting (di)graph analogues of the Laplacian acting on differential forms, as a foundation for later developments. Some known cohomology theories of (di)graphs are similar to the conventional cohomologies for topological spaces, but at the same time also seem to exhibit some different and perhaps puzzling features at first glance. Also, one should ask how these different cohomology theories may be treated in a uniform way. The second purpose of this paper, starting with Section 4, is to use sheaf theory to study (di)graph cohomology theories, with the hope of treating different theories within a single framework. It turns out that there is a Poincaré lemma for the path cohomology of digraphs, which follows from the construction of the CW complex mentioned in the previous paragraph. Our approach here is partly inspired by a recent study of topologies on a graph [7], and our motivation partly lies in the hope that the sheaf theory idea might eventually lead to a much-hoped cohomological proof of the Riemann–Roch theorem for graphs [1].

Acknowledgments

The authors thank CASTS (Center of Advanced Study in Theoretical Sciences) of National Taiwan University, where most of the work was done during their visit. They also thank Fan Chung, Alexander Grigor'yan, and Yong Lin for useful discussions, and thank the referee for useful suggestions.

2. From digraph to CW complex

In this section, we follow the notation in [9], with some modifications that we will mention. Let G be a finite digraph. By a primitive allowed k -path, we mean an ordered sequence of vertices $i_0 i_1 \dots i_k$, such that $i_s i_{s+1}$ is a directed edge in G , for $s = 0, 1, \dots, k - 1$. We say this primitive allowed path is regular, iff all these vertices are different from each other. Note that this regularity condition is more restrictive than the one used in [9, Section 2.3]. There are several reasons we prefer this regularity condition: e.g., with this new condition, the homology groups are now

obviously bounded above, and the Lefschetz fixed-point theorem holds,¹ both of which are not true with the old regularity condition. On the other hand, we will make some remarks about relaxing our regularity condition at the end of this section, to extend the applicability of our construction. We let $A_k(G)$ denote the space of regular allowed k -paths, which, by definition, is the free \mathbb{Z} -module generated by all regular primitive k -paths, and let $\Omega_k(G)$ denote the submodule of ∂ -invariant regular allowed k -paths defined recursively, as in [9, section 3.3]; recall that this means the subspace of $A_k(G)$ consisting of elements whose boundary is an element of $A_{k-1}(G)$. We also use $\Omega(G)$ to denote the direct sum of $\Omega_k(G)$ for all k . When no confusion will arise, we omit G and write A_k and Ω_k . We call k the length of the path. Note that $A_k(G) = 0$ when $k \geq |G|$.

For any $P = \sum_{k=1}^m c_k p_k \in \Omega_k(G)$, where $p_k, k = 1, 2, \dots, m$, are primitive regular allowed paths, we define $w(P) = \sum_{k=1}^m |c_k|$ to be the width of the path P . For each p_k , we define its support to be the subgraph it defines—namely, the minimal subgraph of G , such that p_k is an allowed path in the subgraph. We define the support of P to be the union of the support of each p_k , where c_k is nonzero, and denote it by $Supp(P)$. We say P is minimal iff there do not exist integers $d_k, k = 1, 2, \dots, m$, such that $|c_k - d_k| \leq |c_k|$ and $|d_k| \leq |c_k|$ for each $k = 1, 2, \dots, m$, and $P' = \sum_{k=1}^m d_k p_k \in \Omega_k$, and $w(P') < w(P)$. In this definition, if such a P' exists, we say that P' is strictly smaller than P . Note that $Supp(P') \subset Supp(P)$, and we have also that $P - P' \in \Omega_k$ is strictly smaller than P . Therefore, it is clear that any element in $\Omega_k(G)$ is a linear combination of minimal elements.

Lemma 2.1. *Any minimal path is a linear combination of primitive paths with the same starting and ending vertices.*

Proof. Given any two primitive paths with different starting vertices that both show up in a ∂ -invariant path, if some of their boundary components possibly cancel through a string of other primitive paths, then at certain step one has to change the starting vertex, and therefore the cancellation is not useful in eliminating non- ∂ -invariant paths, in the sense that there exists a strictly smaller ∂ -invariant path consisting of primitive paths all starting with the same vertex. The same argument applies to the ending vertex. \square

Now we are going to construct cells from minimal paths and construct a CW complex given a choice of the integral basis of $\Omega(G)$ consisting of minimal paths, whose existence is a corollary of Lemma 2.3, below, which we will prove together with Lemma 2.2 simultaneously by induction.

Lemma 2.2. *Any minimal path P is a linear combination of primitive paths, with coefficients being either 1 or -1 .*

Lemma 2.3. *Any minimal integral relation among minimal paths of a fixed length is of the form $\sum_{i=1}^m \lambda_i P_i = 0$, where all the coefficients λ_i are either 1 or -1 . Here, the definition of minimal integral relations are the same as that in the definition of minimal paths—in an obvious sense that it cannot be written as a sum of two strictly smaller relations.*

¹We will explain this briefly in Section 5.

Note that Lemma 2.3 implies that any rational basis of $\Omega(G)$ consisting of minimal paths is an integral basis; thus it implies the following:

Corollary 2.4. *There exists an integral basis of $\Omega(G)$ consisting of minimal paths.*

For path length $k = 1$, both lemmas are obviously true, and, furthermore, one associates a k -cell to any minimal path of length k , by filling in a $(k - 1)$ -sphere corresponding to the union of cells associated with boundary components of the path, meaning that the boundary of the path decomposes uniquely as a sum of smaller minimal paths of length $k - 1$, to each of which we have associated a cell, and the cell association commutes with the boundary operation. Now suppose all of these statements are true for path length up to $k - 1$, and take P to be a minimal path of length k . As ∂P is a path, it can be decomposed into a sum of minimal paths of length $k - 1$, where all the paths are smaller than or equal to ∂P , for which there are associated $k - 1$ cells. The union of these cells, counting multiplicity, is a sum of closed manifolds, as $\partial\partial P = 0$. Note: the reader can convince himself/herself that each minimal path of length $k - 2$ that shows up as a boundary component of a boundary component of P appears even times as expected, and in particular there are no singularities on these manifolds. We construct a height function on it as follows: By Lemma 2.1, the starting and ending vertices of any minimal path are unique. First of all, there is a height function on edges, given by piecewise-linearly extending the integer-valued length function defined on vertices, given by the position it sits in a primitive path component (note that this position number is the same for any primitive path one chooses, as a consequence of the obvious fact that any primitive path of maximal length in the support of a minimal path must be a component of the minimal path). We proceed by extending the height function to disks and so on, as we can always extend the height function from a sphere to the ball it bounds. Take any of these closed manifolds and call it M . We can make a small perturbation to make the height function become a Morse function on M . We single out a subset E_1 of the set of vertices in the support of P , consisting of vertices such that any path in the support of P connecting the vertex to the ending vertex E is of length 1. We define another subset S_1 in the symmetric way, with respect to the starting vertex. It is clear that the only possible critical points of this height function are the starting and ending vertices, and vertices in E_1 and S_1 , as aside from them there is always a direction in which the function is strictly monotonic. Now suppose a vertex W in E_1 is a critical point; then E cannot lie on M . Take any primitive path component p of P (whose support after truncating E is in M) that goes through W ,² and let us write it as $p = SqWE$. Then qW as a boundary component of SqW has to be cancelled by a boundary component of a primitive path in a minimal path associated with M , all of which consist of primitive paths of length $k - 1$ that do not go through E . So the only such possible primitive path that has a boundary component cancelling it is itself with a different orientation, which is a contradiction. For the same reason, vertices in S_1 cannot actually be critical points. Therefore, the only possible critical points are S and E , which implies that M is a $(k - 1)$ -sphere and S, E are in its support. Now take all length k primitive paths in the support of M , defined by the union of the support of $k - 1$ minimal paths associated with M , with orientation determined by

²Such a primitive path has to exist in the situation.

orientations of boundary components; then we get a ∂ -invariant k -path that is smaller than or equal to P . Recall that any maximal-length primitive path in the support of P is a primitive component of P . Also note that for any primitive path of length k , all of its boundary components must have support in a single sphere, as otherwise there have to exist boundary components that do not belong to any of the spheres, which is impossible. On the other hand, any primitive $(k - 1)$ -path associated with M must be a boundary component of a primitive k -path with support in M , as a consequence of the fact that any longest primitive path in the support of M must have length k . So the path we just constructed has to be equal to P as P is minimal, and therefore ∂P corresponds to a single $(k - 1)$ -sphere. This implies Lemma 2.2 for P , and that the decomposition of ∂P into a sum of minimal elements is unique. On the other hand, we can fill in the sphere to get a k -cell for P . This cell association clearly commutes with the operation of taking boundary, by construction.

Let us choose an integral basis for each $\Omega_j(G)$ consisting of minimal paths, for j up to $k - 1$, where Lemma 2.3 is true by inductive hypothesis. We now construct a $(k - 1)$ -skeleton together with some k -cells that we will later use in the induction. For this purpose, we need to possibly exclude cells associated with minimal paths that are not elements of the basis chosen, and some cells may need to be modified accordingly.

Again, for any vertex, one associates a zero-cell. For any edge, one associates a one-cell with boundary given by the boundary of the edge.

Suppose again that for all minimal paths in our basis of length up to $i - 1$, ($i \leq k$) one has associated cells of the corresponding dimension (by filling in a sphere that is associated with the boundary of the path) so one has a CW complex with cell dimensions up to $i - 1$. Now pick any designated minimal path P of length i —it must have a single starting vertex S and a single ending vertex E by Lemma 2.1. Again, minimality of P and Lemma 2.2 imply that the boundary of P can be decomposed uniquely into an integral linear combination of minimal paths, with all coefficients being 1 or -1 . For any minimal path P' of length $i - 1$ that shows up in the linear combination, if it is in our chosen basis, we have already assigned a $(i - 1)$ -cell to it. Otherwise, it is a unique integral linear combination of basis elements, with coefficients being 1 or -1 by Lemma 2.3.

If we union the cells in the previous paragraph associated with basis elements in the linear combination, we get a manifold with boundary being an $(i - 2)$ -sphere specified by the union of all $(i - 2)$ -dimensional cells associated with $\partial P'$.³For any such manifold, we can again construct a Morse height function by gluing together individual such functions on cells associated with each minimal element—note that the height function is constructed in a way that enables one to glue. Then the same argument shows that it is an $(i - 1)$ -disk: e.g., one can attach another $(i - 1)$ -cell with the $(i - 2)$ -sphere, to get a manifold without boundary, and then use the same Morse theory argument.

Therefore, for each minimal path that shows up in the above decomposition of the boundary of P , there corresponds a piece of the already existing CW complex homeomorphic to an $(i - 1)$ -disk, with boundary as we described. So again the boundary of these disks cancel, and therefore the union of all of them is a manifold of

³Note that a choice of basis is important for this to be true.

dimension $i - 1$. Note: one has to show further that each $(i - 1)$ -cell associated with basis elements that shows up has multiplicity 1 (or -1), when taking all the P' into account. This can be done by first restricting our attention to all the P' that share the same starting and ending vertices. Each of these P' corresponds to a previously constructed $(i - 1)$ -cell. One sees that the union of these cells is homeomorphic to a $(i - 1)$ -disk, as a consequence of the fact that all these $(i - 1)$ -cells of different starting and ending vertices union to form a manifold without singularity. Then, if any of the $(i - 1)$ -cells aforementioned has greater multiplicity, by an argument similar to that in the inductive proof of Lemma 2.3, below, a sphere must be present, resulting from gluing together cells associated with certain basis elements that show up, so one creates a nontrivial linear relation among basis elements, which is impossible. Now the same Morse height function argument shows that this manifold is homeomorphic to a $(i - 1)$ -sphere, and therefore one can fill it in with an i -cell.

The previous induction goes up to $i = k$. To continue, we have yet to finish our inductive proof of Lemma 2.3 for length k . Suppose we have a minimal integral relation among minimal paths of length k ; then obviously we have unique starting and ending vertices for all primitive paths involved in this relation. So any such integral relation gives rise to a geometric consequence such that the union of all these cells that we have just constructed corresponding to the minimal paths that show up in the relation, counting multiplicity, is a sum of manifolds without boundary.⁴ For any such manifold, we can again construct a Morse height function by gluing together individual such functions on cells associated with each minimal element; then the same argument shows that it is a sphere, and therefore corresponds to a minimal relation, as one easily convinces oneself. So any minimal relation corresponds to a single sphere. Lemma 2.3 is thus evident for k -paths. Our induction is thus complete.

Therefore, we can choose a basis for $\Omega_k(G)$ consisting of minimal paths, and this inductive procedure continues until one associates a cell to each basis element one has chosen, and therefore ends up with a k -skeleton. Now one can simply take k to be the upper bound where $\Omega_k(G)$ is nonzero, and one ends up with a CW complex associated with a choice of integral basis of $\Omega(G)$ consisting of minimal paths. It is evident from the construction that the cell association still commutes with the boundary operator and the integral singular cohomology of the CW complex is canonically isomorphic to the digraph path cohomology.

Our next step is to construct a homotopy between any such CW complexes. For this purpose, we focus on each k step by step, where k is the length of path.

Next, we let a_1, \dots, a_s be any other integral basis of $\Omega_k(G)$ consisting of minimal elements. Then the change of basis from b_1, \dots, b_s to a_1, \dots, a_s can be done in a sequence of s steps, where each step can be expressed in the form $c_1, c_2, \dots, c_s \rightarrow d_1, c_2, \dots, c_s$; this change corresponds to a change of a single basis element from c_1 to d_1 , corresponding to a minimal integral relation expressing d_1 as an integral linear combination involving c_1 of the basis elements c_1, c_2, \dots, c_s . By Lemma 2.3 and its proof, we see there is the following continuous map of topological spaces that we can define: c_1 corresponds to a cell. Write it as the unique integral linear combination of d_1, c_2, \dots, c_s . We “collapse” this c_1 cell onto the union of cells corresponding to

⁴Again, a choice of basis that we have already done up to length $k - 1$ is important for this to be true.

this integral linear combination, which can be viewed as a refinement of the c_1 cell prescribed by this linear relation. This procedure does not affect cells of strictly lower dimensions, and it is clear that this “collapsing” can be extended to a continuous map of the two CW complexes corresponding to these two different basis: i.e., one extends this map in an obvious way to higher dimensional cells. There is of course a continuous map in the reversed direction by collapsing from the second basis to the first basis. One checks directly that the composition of these two maps is homotopic to the identity map, basically by “slowly pulling the string back.” Therefore, any such collapsing is a homotopy.

Therefore, for each digraph G , one assigns a CW complex unique up to homotopy. We next explain that this assignment is functorial, meaning that it defines a functor from the category of digraphs, where morphisms are defined in a particularly strict sense that we will explain below, to a skeleton of the homotopy category of CW complexes. For this homotopy category, we mean that the objects are CW complexes, while the morphisms are homotopy classes of continuous maps of topological spaces. Given any map from G to another digraph G_1 (which means that vertex maps to vertex and directed edge maps to directed edge) that preserves the incidence relations among directed edges and vertices. For our first discussion below, we do not allow different vertices to map to the same vertex, and we only establish the functoriality below in this narrow sense. We will see the discussion can probably be extended in a larger category, where functoriality in a broader sense should hold. It is clear that any minimal path is mapped to a linear combination of allowed ∂ -invariant paths of the same length, which is a sum of minimal paths. Picking any integral basis of $\Omega(G)$ and $\Omega(G_1)$ consisting of minimal paths, and denoting by T_G and T_{G_1} the corresponding CW complexes, we construct a continuous map from T_G to T_{G_1} inductively: First, vertices and directed edges are mapped to their images. Now suppose cells corresponding to minimal paths of length strictly less than k are mapped; then for any minimal path P of length k in the chosen basis, the image can be decomposed as a sum of minimal paths of G_1 , again with coefficients being 1 or -1 , which themselves then correspond to unions of k -cells in T_{G_1} homeomorphic to k -disks with boundary corresponding to the boundary of the minimal paths, and furthermore any basis element that appears is with multiplicity 1 or -1 . Thus one can homeomorphically map the k -cell associated to P , to the union of these k -cells, in terms of a refinement (subdivision) of the cell, which gives the desired map inductively. Note that this procedure does not affect maps of cells of strictly lower dimensions that are already defined. It is then routine to check the functorial properties, as refinements compose in a desired way.

When directed edges are allowed to collapse, and, in particular, different vertices are allowed to map to the same vertex, a digraph may be mapped to a multi-digraph, which means multi-edges with arbitrary orientations and self-loops are allowed.⁵ In the larger category of multi-digraphs, a morphism is defined to be a map that takes vertices to vertices, and directed edges to directed edges, such that it preserves the incidence relations among directed edges and vertices. To extend our discussion to this larger category, we need to relax our definition of a path and the regularity condition,

⁵It can also happen that a digraph still maps to a digraph, but our previous discussion may encounter problems of degeneration.

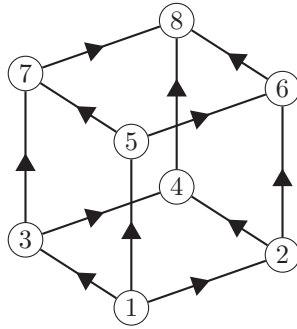


Figure 1: The CW complex of this digraph is a cube.

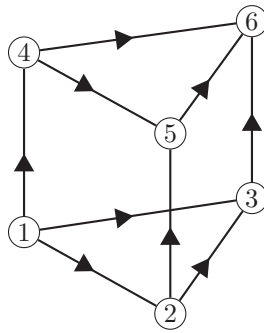


Figure 2: The CW complex of this digraph is a triangular prism.

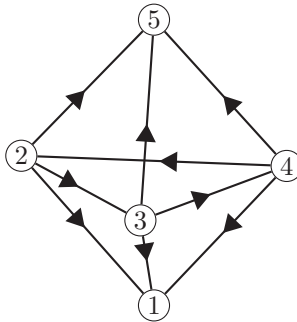


Figure 3: The CW complex of this digraph is a 2-sphere.

in a precise way that allows paths that result from various kinds of degenerations. We exhibit in the following a candidate choice of such definitions.

Definition 2.5. A virtual primitive path is an ordered string of vertices $V_0V_1 \cdots V_s$, together with the following data: for any pair of consecutive vertices V_k, V_{k+1} in the string ($k = 0, 1, \dots, s-1$), either one specifies a directed edge connecting them or V_kV_{k+1} is not a directed edge, and, furthermore, for any consecutive pairs of vertices

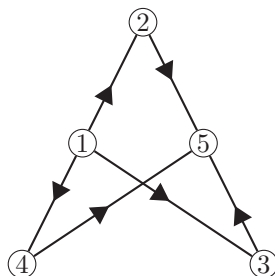


Figure 4: Choosing any integral basis of $\Omega_2(G)$ consisting of minimal 2-paths, the CW complex of this digraph is formed by two 2-cells glued along two 1-cells via an identification, and it is homotopic to a point. Note that if one naively constructs the CW complex by associating a cell to each minimal path, one would get a 2-sphere instead, which would have a different cohomology than the path cohomology.

in any string of vertices that appear as a (formal) component of $\partial(P)$, or boundary components of boundary components, and so on, one either specifies a directed edge connecting them or there is no directed edge between them. These specifications have to be done in a way compatible with all incidence relations among paths.

Definition 2.6. A primitive path is a virtual primitive path, such that for any pair of consecutive vertices V_k, V_{k+1} in the string ($k = 0, 1, \dots, s - 1$), either we specified a directed edge connecting them or $V_k V_{k+1}$ is not a directed edge but $V_k = V_{k+1}$.

Definition 2.7. If a primitive path satisfies the additional condition that any directed edge appears at most once as a segment of the path and any virtual primitive path of length 1 less appears at most once in the formal components of $\partial(P)$ before any cancellation, then we call it a regular primitive path.

Definition 2.8. A path is an integral linear combination of regular primitive paths, and a ∂ -invariant path is a path, whose ∂ is a linear combination of primitive paths. The space of ∂ -invariant paths is denoted by Ω .

Remark 2.9. This definition reflects the fact that it is possible that some boundary components of a cell collapse while the cell itself stays a cell. So we do not require all boundary components to be regular.

Now we can formulate the following:

Question 2.10. Generalize the construction of the CW complex to multi-digraphs, so that functoriality holds in the broader sense stated. One may need to define the cohomology with a little more care, similar to what is done in [9], to account for the new regularity condition. The cells in this more general setting could probably be regarded as obtained by certain contractions from the cells in the old setting. For functoriality, given two multi-digraphs G, G_1 , a morphism between them, and a minimal path P in G , one in general may need to contract the cells associated with P in the way prescribed by the multi-digraph morphism and then do the map described above to match the choice of integral basis of the second multi-digraph. Note that a cell may be mapped to lower-dimensional cells in general.

Remark 2.11. Note that for the category of digraphs with morphism defined in our narrow sense the resulting CW complex has the property that any attaching map is an obvious homeomorphism. However, things will be more complicated in the bigger category of multi-digraphs.

Returning to digraphs, by Chapter 7 of [9] it is evident that given an integral basis for two digraphs, their product is an integral basis of the product digraph, and furthermore that taking boundaries of products of paths satisfies the Leibniz rule, which implies that our association of a CW complex to a digraph preserves products.

Remark 2.12. As we have seen, one can associate a cell to any minimal path, and thus actually construct a CW complex in a canonical way from G without a choice of integral basis as above, and the construction also probably will have all these nice functorial properties. However, the cohomology of this new CW complex will get additional contributions from linear relations among minimal paths, which perhaps makes this construction less appealing.

3. Some immediate consequences

It then follows from standard algebraic topology that a homotopy between digraphs induces isomorphisms of cohomology groups, and that the Künneth formula holds for digraph cohomology. Furthermore, one can define arbitrary higher homotopy groups of a digraph in terms of that of the CW complex. On the other hand, the cohomology of digraphs becomes a functorial graded ring as that of the CW complex is such a graded ring under the cup product. It turns out that this product can be lifted to the level of forms, which are defined to be elements in Ω^k , the dual of Ω_k , and the lift still respects associativity and the Leibniz rule and is functorial. Most of these are first proved in [9, 10]. We show below that a very simple formula exists for this lifted product,⁶ which may be relevant, for example, in studying some gauge field theories on the digraph.

One sees from the construction of the CW complex that, for any minimal path in the chosen basis $P = \sum_{k=1}^m c_k p_k \in \Omega_k(G)$, where $p_k, k = 1, 2, \dots, m$, are primitive regular allowed paths, there exists a unique subdivision of cells, given by adding edges to connect all previously non-neighboring pairs of vertices in every p_k in the same direction of the path so that each p_k becomes a complete graph. After this subdivision, the cell associated with P is divided into a sum of simplexes, each associated with a p_k with the newly connected edges. One can do this subdivision to all cells associated with basis elements in a consistent way, and then the CW complex becomes a simplicial complex whose simplicial cohomology is canonically isomorphic to the singular cohomology of the CW complex. The cup product in this simplicial complex has the well-known simple formula in terms of simplexes, which then translates into the corresponding formula for the cup product in the CW complex restricted to the cells we are considering. Unravelling the definitions, one sees that this restriction actually gives rise to the formula for the functorial cup product of forms on digraphs. Let $\alpha \in \Omega^p(G)$, and $\beta \in \Omega^q(G)$, and $k = p + q$. Suppose $p_k = V_0 \cdots V_{p+q}$. We let $p_k|_{0 \dots p}$

⁶The formula is actually implicitly contained in [10], but here we provide a more geometric understanding of it.

and $p_k|_{p\dots p+q}$ denote the allowed paths $V_0 \cdots V_p$ and $V_p \cdots V_{p+q}$, respectively, resulting from truncating p_k in the way described. Then we have the formula for the cup product $\alpha \cup \beta$ on P as follows:

$$\alpha \cup \beta(P) = \sum_{k=1}^m c_k \alpha(p_k|_{0\dots p}) \beta(p_k|_{p\dots p+q}).$$

Note that the above formula does not make sense at a first glance, as each individual truncation may not be in Ω ; however, the formula is understood in the sense that one needs to first merge together all terms with the same truncation in the arguments⁷ in the above sum. Then it is an easy exercise to show that it indeed makes sense after the merging. One sees also from this formula that it is independent of our choice of basis.

Regarding the homotopy invariance property, note that a homotopy of two maps of digraphs is defined in direct analogy with the corresponding concept in topology [10], and applying our functor, any such homotopy gives rise to a homotopy between two continuous maps of the CW complexes associated with the two digraphs, therefore inducing isomorphic maps on cohomology groups. Furthermore, since homotopy of digraphs becomes homotopy of the CW complexes, the functor provides a tool to study homotopy properties of digraphs.

4. Clique cohomology

In this section, we start to use sheaf theory to reformulate some known (di)graph cohomology theories. We first illustrate the ideas with the example of clique cohomology, and here we try to follow notations in [6].

Let G be a finite graph.⁸ Let \mathcal{G}_k denote the set of all K_{k+1} subgraphs of G , and $\mathcal{G} = \cup_{k=0}^\infty \mathcal{G}_k$. By a topology T on a graph G , we actually mean a topology T on the set \mathcal{G} . Given any topology, one can consider the category of sheaves of abelian groups on \mathcal{G} . Sheaf cohomology is well defined, as any such category has enough injectives. However, it is crucial that one chooses a suitable topology for all applications that follow. To mimic the case of usual continuous geometry, here we consider the unit ball topology, which is defined by a topology subbasis as the set of all unit balls, whose definition we state below:⁹

For any vertex $v \in G$, we define its unit ball subgraph B_v as the subgraph of G generated by v and all of its neighbors. In other words, it is the largest subgraph of G containing only these vertices. For each B_v , we canonically associate a subset \mathcal{B}_v of \mathcal{G} as follows: $x \in \mathcal{G}$ lives in \mathcal{B}_v if and only if x is a subgraph of B_v .

It is clear from definition that these \mathcal{B}_v give a subbasis of topology. Note that B_v is a cone; therefore, one has $H^i(B_v) = 0$, for all $i > 0$ —here H^i is the graph cohomology functor defined by the clique complex. For any $x \in \mathcal{G}$, let us denote by \mathcal{U}_x the smallest open subset containing x , which always exists as there are finitely many such open sets. By our choice of topology, \mathcal{U}_x corresponds uniquely to a subgraph U_x in the same

⁷Namely, the same $\alpha(p_k|_{0\dots p})$ or $\beta(p_k|_{p\dots p+q})$.

⁸More generally, the following theory also works for an infinite graph all of whose vertices have finite degree.

⁹Note there probably exist other good choices of topology for our purpose here.

sense that \mathcal{B}_v corresponds to $B_v - y \in \mathcal{G}$ lives in \mathcal{U}_x if and only if y is a subgraph of U_x . We have the following characterization of U_x :

Lemma 4.1. U_x is the intersection of maximal complete subgraphs containing x .

Proof. Suppose a vertex v does not belong to some maximal complete subgraph K containing x . Then there exists a vertex w in K such that w is not connected to v by an edge. Then B_w contains U_x , but, on the other hand, v is not in B_w . So v is not in U_x , which implies that $U_x \subset K$, and so U_x is contained in the intersection of maximal complete subgraphs $\cap K$ containing x . Conversely, if v_1 is a vertex such that B_{v_1} contains x , then the complete graph K_1 containing both v_1 and x is a subgraph of G . Consider the maximal complete subgraph K_2 of G containing K_1 : we have that $K_2 \subset B_{v_1}$ by the definition of the unit ball subgraph. So $\cap K \subset K_2 \subset B_{v_1}$, which proves the inclusion in the other direction. \square

As a consequence, we have the following

Corollary 4.2. U_x is a complete subgraph, and, in particular, $H^i(U_x) = 0$ for any $i > 0$.

Take A to be any abelian group; next, we construct a flasque resolution of the constant sheaf \mathcal{A} on \mathcal{G} with values in A .

Take $\mathcal{U} \subset \mathcal{G}$ an open subset. Let \mathcal{U}_k denote $\mathcal{G}_k \cap \mathcal{U}$. Define $\mathcal{C}^k(\mathcal{U})$ to be the abelian group of continuous functions, from \mathcal{U}_k to A , where \mathcal{U}_k is equipped with the subset topology (which actually does not matter) and A the trivial topology. It is easy to check from the definition that the assignment $\mathcal{U} \rightarrow \mathcal{C}^k(\mathcal{U})$ defines a flasque sheaf \mathcal{C}^k on \mathcal{G} . The differential of the clique complex gives rise to a differential mapping \mathcal{C}^k to \mathcal{C}^{k+1} , making it into a complex of sheaves. Furthermore, any section of the constant sheaf \mathcal{A} on \mathcal{U} is a function that is constant on every connected component of \mathcal{U} and thus can be mapped to a section of $\mathcal{C}^0(\mathcal{U})$ by associating the vertices in each connected component with the corresponding constant value in A . We have the following:

Lemma 4.3. \mathcal{C}^k gives a flasque resolution of the constant sheaf.

Proof. The exactness at \mathcal{A} and \mathcal{C}^0 is obvious. At general \mathcal{C}^k , we look at each stalk. Unraveling the definition, the exactness after taking stalks reduces to Corollary 4.2. \square

Taking global sections, we therefore have the following:

Theorem 4.4. There is a canonical isomorphism $H^i(\mathcal{G}, \mathcal{A}) \cong H^i(G, A)$, where $H^i(G, A)$ denotes the graph cohomology defined by cliques, taking values in A .

Next, we consider Čech cohomology. Taking a finite open cover $\mathcal{U}_i, i = 1, 2, \dots, s$, of \mathcal{G} , one forms the Čech complex for any sheaf \mathcal{F} of abelian groups. As will be expected, we have the following:

Lemma 4.5. For each i , there is a natural map $\check{H}^i(\mathcal{G}, \mathcal{F}) \rightarrow H^i(G, \mathcal{F})$, functorial in \mathcal{F} .

Proof. See [5, III.4]. \square

Take $\mathcal{F} = \mathcal{A}$. Take a finite open cover $\mathcal{U}_i, i = 1, 2, \dots, s$, of \mathcal{G} such that any intersection has trivial higher cohomology. We as usual have the following:

Theorem 4.6. *The natural map above gives an isomorphism $\check{H}^i(\mathcal{G}, \mathcal{A}) \cong H^i(\mathcal{G}, \mathcal{A})$.*

Proof. See proof of Theorem 4.5 on page 222 of [5]. \square

Remark 4.7. Note that such an open covering always exists, and 4.1 provides a canonical one as such, by 4.2 and 4.4 and the evident fact that the subset topology on any open set corresponding to a subgraph coincides with the unit ball topology of the subgraph.

Remark 4.8. Let us look at how the graph cohomology may be glued from smaller pieces, at least in principle. Let \mathcal{U} be any open subset of \mathcal{G} corresponding to a subgraph U . Let us denote the closed subset $\mathcal{Y} = \mathcal{G} - \mathcal{U}$. Then all statements of exercises 2.3 and 2.4 on page 212 of [5] apply. In particular, taking $\mathcal{F} = k$, we get the following long exact sequence:

$$0 \rightarrow H_{\mathcal{Y}}^0(\mathcal{G}, k) \rightarrow H^0(\mathcal{G}, k) \rightarrow H^0(\mathcal{U}, k) \rightarrow H_{\mathcal{Y}}^1(\mathcal{G}, k) \rightarrow \dots,$$

where $H^i(\mathcal{G}, k)$ and $H^i(\mathcal{U}, k)$ are naturally isomorphic to the usual graph cohomology, as we have seen. The additional piece $H_{\mathcal{Y}}^i(\mathcal{G}, k)$ may be analyzed by the same flasque resolution 4.3. Furthermore, this cohomology with support in \mathcal{Y} satisfies the excision and Mayer–Vietoris sequence. A tricky thing is that the combinatorial translation of such statements may not be nice or very useful in general.

5. Some comments

If one regards a graph G simply as a one-dimensional simplicial complex and considers its simplicial cohomology, the procedure can again be discretized in the same way: one takes the set $\mathcal{G}' = \mathcal{G}_0 \cup \mathcal{G}_1$ and takes all the star graphs [7] as the subbases of topology; then, in the same way, one can show that the cohomology of the constant sheaf realizes this trivial version of graph cohomology.

Next, let us take a look at a simple version of the Lefschetz fixed-point theorem for graphs [8],¹⁰ which states that for any automorphism f of a graph G , one defines its Lefschetz number as

$$\Lambda(f) = \sum_{i=0}^{\infty} (-1)^i \text{Tr}(f^* : H^i(G, k) \rightarrow H^i(G, k)).$$

Then if $\Lambda(f)$ is nonzero, f has at least one fixed simplex, where k is any ground field and $H^i(G, k)$ is graph cohomology taking values in k . The proof of this can be reduced to the familiar case of simplicial complexes, or one shows as usual that it is a consequence of linear algebra.

From our framework, for any injective graph homomorphism $\phi : G_1 \rightarrow G_2$, since it maps cliques to cliques, one has an induced continuous map of topological spaces $\mathcal{G}_1 \rightarrow \mathcal{G}_2$, which we still denote by ϕ if no confusion arises. In particular, if $\phi = f$

¹⁰One can also consider more elaborate versions, but here we take the simplest version for the purpose of illustration.

is an automorphism of G , f is continuous as a map from \mathcal{G} to itself. On the other hand, by 4.4, the definition of $\Lambda(f)$ can also be stated using $H^i(\mathcal{G}, k)$. Therefore, the above Lefschetz theorem for graphs is equivalent to a Lefschetz theorem for f and the topological space \mathcal{G} with Lefschetz number defined by the sheaf cohomology. It looks to be an interesting question to elaborate on this observation, from the point of view of finite set topology.

Also one notes that the same theorem holds for digraphs: with regard to Lemma 2.1, a morphism of digraphs that has no fixed vertices has to have zero trace in Ω_k , and thus its Lefschetz number has to be zero.

6. Path cohomology

For all k , choose any integral basis of $\Omega_k(G)$ consisting of minimal elements, and let us call it B_k . Let X_G denote the union of these basis elements as a set. For any path $P \in X_G$ of length k , we denote by G_P the smallest subgraph of G , such that $P \in \Omega_k(G_P)$. For each P , we define U_P inductively to be the union of $\{P\}$ and U_Q , where Q is any element in X_G that appears as a direct summand of an element in $\Omega_k(G_P)$. We define a topology T on X_G by claiming all U_P to form a subbasis of topology. We then have the following:

Lemma 6.1. U_P is the smallest open subset containing P , and $U_{P_1} \cap U_{P_2} = \cup_{x \in U_{P_1} \cap U_{P_2}} U_x$.

Proof. Check by the definitions. □

Thus we have:

Corollary 6.2. U_P form a basis of topology.

Remark 6.3. The definition of X_G and U_P is carefully chosen so as to take into account the subtle issues involved in the definition of the path cohomology.

For any k , we define a sheaf \mathcal{C}^k of abelian groups on the topological space X_G as follows: for any open set U , one assigns the abelian group of integer valued \mathbb{Z} -linear functions on the \mathbb{Z} -module spanned by the set of length k elements in U . It is obvious that \mathcal{C}^k is a flasque sheaf. It is straightforward to check that the sheaves \mathcal{C}^k form a complex of sheaves via the natural differential. Therefore, taking global sections, the cohomology of this complex of sheaves computes the path cohomology of digraphs. We have the following lemma.

Lemma 6.4 (Poincare lemma). \mathcal{C}^k is a flasque resolution of the constant sheaf.

And a simpler version:

Lemma 6.5 (Poincare lemma: baby version). For any $P \in X_G$, we have $H^i(G_P) = 0$ for all $i > 0$.

Proof. The combinatorics of both lemmas are subtle, and the authors only know a combinatorial proof of the baby version lemma. On the other hand, unravelling the definitions, they evidently follow from the construction of the CW complex in

Section 2—namely, it follows from the proof that in the inductive process of constructing the CW complex, or in associating a cell to any minimal path, ∂P gives rise to a $(k - 1)$ -sphere, for P a minimal k -path. \square

Thus we have:

Theorem 6.6. *The cohomology of the constant sheaf on X_G is naturally isomorphic to the path cohomology of G .*

7. Computation: a first discussion

In this section, we are concerned with the computation of the CW complex and the cohomology, and try to get some first understanding of the complexity. We have the following:

Lemma 7.1. *For digraphs with a uniform bound on the vertex degree, if one fixes k , then the time complexity of computing a basis of Ω_k consisting of minimal paths, and thus the k skeleton of the CW complex, is quadratic.*

Proof. Let D denote the uniform bound of vertex degree, and let n be the number of vertices of the digraph. By Lemma 2.1, any minimal path has unique starting and ending vertices. There are at most $n(n - 1)$ choices of these ordered pairs of vertices. For each such choice, there are at most D^{k-1} many primitive paths of length k with the given starting and ending vertices, and once all these primitive paths are enumerated, one is left with another finite calculation to determine a rational basis of minimal paths with given starting and ending vertices (think of Lemma 2.2). These basis elements combine to give a desired basis of Ω_k consisting of minimal paths. \square

Remark 7.2. The proof that the homotopy type of the CW complex is determined by the digraph obviously also shows that the same is true for any k skeleton.

In the following, we present a recursive scheme for computing a basis of Ω_k consisting of minimal paths.

Take any minimal path P of length k , and with starting vertex S and ending vertex E . As before, we single out a subset E_1 of the set of vertices in the support of P , consisting of vertices such that any path in the support of P connecting the vertex to E is of length 1. It is then clear that, for any vertex W_1 in E_1 , if one groups together all primitive paths in P going through W_1 taking signs into account and truncates E from them, then one gets a path P' of length $k - 1$, and, furthermore, P' is ∂ -invariant: the proof of this is essentially the same as that of Lemma 2.1. So P' can be written uniquely as a sum of basis elements of length $k - 1$ that has already been computed, again with all the coefficients being either 1 or -1 , and, furthermore, the union of these $(k - 1)$ -cells corresponding to the basis elements that show up is homeomorphic to a $(k - 1)$ -disk. (See previous arguments in constructing the CW complex.) For P' , one again defines a set E_2 to be the subset of vertices in the support of P' such that any path connecting the vertex to W_1 is of length 1. One then sees that in order for P to be ∂ -invariant, it is necessary and sufficient that for any vertex W_2 in E_2 that is not connected to E by a directed edge and any primitive path P'' in P' that goes through W_2 , there exists another vertex W in E_1 such that, once

one expresses the same truncation of (signed) summation of all primitive paths of P that goes through W in terms of the unique linear combination of length $k - 1$ basis elements chosen, there exists one basis element in the linear combination that contains a primitive path given by switching the ending vertex of P'' from W_1 to W , with appropriate sign, so that boundary components of these two primitive paths given by deleting W_1 and W cancel as desired. In this way, one finds all ∂ -invariant paths of length k between S and E , and then one goes on to find the minimal ones, and a rational thus integral basis, for which efficient and straightforward algorithms exist.

Remark 7.3. It is clear that the above recursive scheme will be more efficient than a basic brute force algorithm, following from the proof of Lemma 7.1. It is a problem to carefully study the complexity of such an algorithm in more general situations.

References

- [1] M. Baker, and S. Norine, Riemann-Roch and Abel-Jacobi theory on a finite graph, arXiv:math/0608360.
- [2] B. Chen, S-T Yau, and Y-N Yeh, Graph homotopy and Graham homotopy, *Discrete Math.* **241** (2001) 153–170.
- [3] D. Happel, Hochschild cohomology of finite dimensional algebras, *Lecture Notes in Math.* **1404**, Springer-Verlag 1989. 108–126.
- [4] J. Friedman, Sheaves on graphs, their homological invariants, and a proof of the Hanna Neumann conjecture, arXiv:1105.0129.
- [5] R. Hartshorne, *Algebraic Geometry*, GTM 52, Springer-Verlag, 1977.
- [6] O. Knill, The Dirac operator of a graph, arXiv:1306.2166.
- [7] O. Knill, A notion of graph homeomorphism, arXiv:1401.2819.
- [8] O. Knill, A Brouwer fixed point theorem for graph endomorphisms, arXiv:1206.0782.
- [9] A. Grigor'yan, Y. Lin, Y. Muranov, and S-T Yau, Homologies of path complexes and digraphs, arXiv:1207.2834.
- [10] A. Grigor'yan, Y. Lin, Y. Muranov, and S-T Yau, Homotopy theory for digraphs, arXiv:1407.0234.
- [11] Z-C. Gu, F. Verstraete, and X-G. Wen, Grassmann tensor network states and its renormalization for strongly correlated fermionic and bosonic states, arXiv:1004.2563.

An Huang anhuang@math.harvard.edu

Department of Mathematics, Harvard University, Cambridge, MA 02138, USA.

Shing-Tung Yau yau@math.harvard.edu

Department of Mathematics, Harvard University, Cambridge, MA 02138, USA.