HARMONIC REPRESENTATIVES IN HOMOLOGY OVER ARBITRARY FIELDS

MICHAEL J. CATANZARO AND BRANTLEY VOSE

ABSTRACT. We introduce a notion of harmonic chain for chain complexes over fields of positive characteristic. A list of conditions for when a Hodge decomposition theorem holds in this setting is given and we apply this theory to finite CW complexes. An explicit construction of the harmonic chain within a homology class is described when applicable. We show how the coefficients of usual discrete harmonic chains due to Eckmann can be reduced to localizations of the integers, allowing us to compare classical harmonicity with the notion introduced here. We focus on applications throughout, including CW decompositions of orientable surfaces and examples of spaces arising from sampled data sets.

1. INTRODUCTION

The classical Hodge decomposition theorem [14, 28] states that the set of all differential k-forms on a smooth, compact, Riemannian manifold M decomposes as a direct sum

(1)
$$\Omega^{k}(M) \cong \mathcal{H}^{k}(M) \oplus \mathcal{H}^{k}(M)^{\perp}$$
$$\cong \mathcal{H}^{k}(M) \oplus \operatorname{im}(d) \oplus \operatorname{im}(d^{*})$$

where we identify $\mathcal{H}^k(M)$ with the space of harmonic k-forms, and d is the exterior derivative, so that $\operatorname{im}(d)$ is the space of exact k-forms and $\operatorname{im}(d^*)$ is the space of co-exact k-forms. We can further identify $\mathcal{H}^k(M)$ with $\ker(d) \cap \ker(d^*)$, as well as with the kernel of the Laplacian $\mathcal{L} = dd^* + d^*d$. The various guises of harmonic forms have created unifying bridges between disparate areas of mathematics. Applications of harmonicity vary from studying heat flow [23, 28] to analyzing random walks [29, 19] and more generally spectral geometry [16, 8], to name a few.

In the 1940s, Eckmann generalized the classical story by replacing the manifold M with a simplicial complex X, the k-forms with the simplicial k-chains, and the Laplacian with an analogously defined discrete Laplacian [11]. In this discrete setting, Eckmann found that a theorem analogous to the Hodge decomposition, sometimes called the *discrete Hodge decomposition*, still holds. That is, the space of k-chains can be written as an direct sum

(2)
$$C_k(X;\mathbb{C}) \cong \ker(\mathcal{L}_k) \oplus \operatorname{im}(\partial) \oplus \operatorname{im}(\partial^*)$$

where \mathcal{L}_k here is the discrete Laplacian $\mathcal{L}_k = \partial \partial^* + \partial^* \partial$. Besides its theoretical significance, this result furnishes many practical applications of harmonicity including voting and ranking theory [9, 25, 15], network theory [1, 2] and spectral graph theory [18].

One particular use of the Hodge decomposition is the following. Every chain in $\ker(\mathcal{L}_k)$ is a cycle, and the natural map $\ker(\mathcal{L}_k) \to H_k(X;\mathbb{C})$ sending each cycle to its homology class is an isomorphism. Inverting this isomorphism lets us select a canonical representative of each homology class. In other words, the Hodge decomposition implies that every homology class has a unique harmonic representative. Among other implications, this result lets allows us to perform homology computations in a subspace instead of a quotient space, the latter being tricky to implement in practice.

The story we have laid out so far holds just as well if we replace \mathbb{C} with any field of characteristic 0. However, we are often interested in homology and cohomology over more general coefficient rings, including fields of positive characteristic. This is especially true in applications; computations over finite fields, especially \mathbb{F}_2 , can be performed more quickly than over \mathbb{Q} , and with perfect precision. One example is the software library Ripser, which is a popular choice for computing persistent (co)homology

[4]. Once we learn of the practical uses of the Hodge decomposition and harmonic representatives, it is natural to ask whether similar results hold over other fields. This leads us to the driving questions of this paper.

- (Q1) Does an analog of the discrete Hodge decomposition exist when we replace \mathbb{C} with a finite field, or more generally with an arbitrary field \mathbb{F} ?
- (Q2) If not, when does every homology class have a unique harmonic representative?

The first of our primary results, Theorem A, provides a list of necessary and sufficient conditions for every homology class in $H_k(X; \mathbb{F})$ to have a unique harmonic representative, addressing (Q2). Subsequently, Theorem B gives stronger necessary and sufficient conditions for the existence of a Hodge decomposition analogous to Eq. (1), namely, if and only if every homology class and every cohomology class have a unique harmonic representative. This answers question (Q1). Sections 3 and 4 are spent proving these results and developing methods for computing these unique harmonic representatives. While our motivating examples arise from CW complexes with finitely many cells in each dimension, in these sections we work in the more general setting of chain complexes of finitedimensional spaces equipped with standard bases. For this reason, our results can be applies just as well to the dual chain complex to derive results about cohomology that are "dual" to our results about homology.

We then move on to consider (Q1) and (Q2) from a different perspective. Given a fixed CW complex X with finitely many cells in each dimension, in Section 5 we explore for which primes p the cellular homology of X over \mathbb{F}_p satisfies the statements of Theorem A. That is, given X, we provide sufficient conditions for p under which every cellular homology class of X over \mathbb{F}_p has a unique harmonic representative. We accomplish this by leveraging the cleaner rational case and descending to the finite field. This method is thwarted if $H_k(X)$ has p-torsion, or if p is one of finitely many primes determined by the combinatorial structure of X. Theorem D determines this list of problematic primes in terms of an integer that we call Υ which depends on the combinatorial structure of X.

In the special case that X is a CW structure on an orientable surface, the statements of Section 5 can be simplified. In Section 6 we show that, in this special case, Theorem C has a simpler statement in terms of the discrete Laplacian, giving us a tidy condition in Theorem 6.9 to ensure the existence and uniqueness of all harmonic representatives on X over \mathbb{F}_p .

Ultimately the difference between harmonicity in positive and zero characteristic stems from the fact that vector spaces over fields of positive characteristic cannot be equipped with norms or inner products, since fields of positive characteristic cannot be ordered. The best replacement is a nondegenerate bilinear form, for which there can exist degenerate subspaces. Degenerate subspaces have nontrivial intersection with their 'orthogonal complement', so a direct sum decomposition as in Eq. (1) is generally not possible. Over a field of characteristic 0, $\ker(\mathcal{L}_k) = \ker(\partial) \cap \ker(\partial^*)$. However, this fact depends on the assumption that every subspace is nondegenerate. Over a field of positive characteristic, we instead have $\ker(\mathcal{L}_k) \supseteq \ker(\partial) \cap \ker(\partial^*)$, often with strict containment. Indeed, over a general field, an element of $\ker(\mathcal{L}_k)$ need not even be a cycle (see Example 3.2). Since this paper is interested in harmonic representatives of (co)homology classes, we make the choice to define the harmonic chains to be the space $\ker(\partial) \cap \ker(\partial^*)$, i.e. the space of chains that are both cycles and cocycles. We emphasize that, over characteristic 0, this is equivalent to the usual definition that instead chooses $\ker(\mathcal{L}_k)$, but in general the two spaces are distinct.

1.1. Related Work. In [10], Ebli and Spreemann utilize the discrete laplacian over \mathbb{R} to develop a homology-sensitive method to cluster the simplices of a simplicial complex. Among other uses, this process produces visualizations of homology classes similar to those in Section 7. A generalization of the discrete Laplacian designed for the setting of persistent homology is introduced and explored in [27] and [20]. In the recent preprint [3], Basu and Cox apply the theory of harmonic representatives to persistent homology to associate concrete subspaces of chain spaces to each bar of the persistence barcode in a stable way. Kališnik et. al. [17] explore related methods for creating a homologically persistent skeleton using higher spanning trees, and encode important topological features of data using these subcomplexes.

2. Background and Main Results

In this section, we briefly recall some facts regarding linear algebra over arbitrary fields, homology, cohomology, and CW complexes that will be used throughout. With this background in place, we then state our main results.

Linear algebra over arbitrary fields. We remind the reader of some basic facts about vector spaces over fields of positive characteristic. A standard reference for this material is given in [24].

A symmetric bilinear form on a vector space V over a field \mathbb{F} is a map $b: V \times V \to \mathbb{F}$ such that

$$b(v, w) = b(w, v) ,$$

$$b(u + v, w) = b(u, w) + b(v, w) , \text{ and}$$

$$b(\lambda v, w) = \lambda b(v, w) ,$$

for any $u, v, w \in V$ and $\lambda \in \mathbb{F}$. Given a symmetric bilinear form b on V and a subset $U \subseteq V$, we define the *orthogonal complement* of U by

$$U^{\perp} = \{ v \in V \, | \, b(u, v) = 0 \text{ for all } u \in U \} \,.$$

The symmetric bilinear form b is said to be non-degenerate if $V^{\perp} = 0$. If U is a subspace of V, then $b|_U: U \times U \to F$ is again a symmetric bilinear form. A subspace U is non-degenerate if $b|_{U\times U}$ is non-degenerate, i.e., $U \cap U^{\perp} = 0$.

Example 2.1. Let V be an n-dimensional vector space over the real numbers $\mathbb{F} = \mathbb{R}$ with basis $\{e_1, e_2, \ldots, e_n\}$. Then $b(e_i, e_j) = \delta_{ij}$, given by the Kronecker delta function, defines a non-degenerate symmetric bilinear form, the standard inner product. Every subspace of V is non-degenerate.

Let W be an m-dimensional vector space over the field with p elements $\mathbb{F} = \mathbb{F}_p$ with basis $\{f_1, f_2, \ldots, f_m\}$. Then $b(f_i, f_j) = \delta_{ij}$ defines a non-degenerate symmetric bilinear form. In this case, b does not define an inner product and there are numerous degenerate subspaces of W, e.g., the one-dimensional subspace spanned by $f_1 + \cdots + f_p$ for $p \leq m$.

2.1. Homology and Cohomology. Let R be a ring. A chain complex C_{\bullet} over R is a sequence of R-modules $\{C_k\}_{k\in\mathbb{Z}}$ called the chain modules, along with homomorphisms $\partial_k : C_k \to C_{k-1}$, called the differentials, satisfying $\partial_{k-1}\partial_k = 0$. The k-cycles $Z_k \subseteq C_k$ are given by the kernels of the differentials $Z_k = \ker \partial_k$. The k-boundaries $B_k \subseteq C_k$ are given by the images of the differentials $B_k = \operatorname{im} \partial_{k+1}$. Finally, since the condition $\partial_{k-1}\partial_k = 0$ implies $B_k \subseteq Z_k$, we can define the homology of the chain complex in degree k to be the quotient $H_k = Z_k/B_k$. We use the notation $[z] \in H_k$ to denote the equivalence class of $z \in Z_k$, and say that z is a representative of [z].

The associated *cochain complex* is the *R*-linear dual of the chain complex. Precisely it is the sequence of *R*-modules $C^k := \operatorname{Hom}_R(C_k, R)$, the *R*-linear dual of C_k , together with codifferentials given by the duals of the original differentials $\partial_k^* : C^{k-1} \to C^k$. The condition $\partial_{k-1}\partial_k = 0$ implies $\partial_k^*\partial_{k-1}^* = 0$. Analogous to cycles and boundaries, the *k*-cocycles $Z^k \subseteq C^k$ are the cochains in the kernel of the codifferential $Z^k = \ker(\partial_{k+1}^*)$, the *k*-coboundaries $B^k \subseteq C^k$ are those in the images $B^k = \operatorname{im}(\partial_k^*)$. and the cohomology of the cochain complex is given by the quotient $H^k = Z^k/B^k$.

Remark 2.2. We are interested in the case where the underlying ring of the chain complex is a field \mathbb{F} , and each chain module C_k is a finite-dimensional vector space with a chosen basis. This is the case when C_{\bullet} arises as a simplicial or cellular chain complex as discussed below in section 2.1 The chosen basis, say $\{\alpha_i\}$ on C_k , defines a symmetric bilinear form b_k on C_k by setting $b_k(\alpha_i, \alpha_j) = \delta_{ij}$, the Kronecker pairing. The form b_k gives rise to a canonical isomorphism between C_k and its dual C^k , letting us identify the two and blur the distinction between a chain c and its dual c^* . This identification is common in the literature, allowing one to think of the codifferentials ∂_k^* as linear maps $C_{k-1} \to C_k$, and think of Z^k and B^k as subspaces of C_k .

Given the identification in Remark 2.2, we state our notion of F-harmonicity.

Definition 2.3. A chain $c \in C_k$ is called \mathbb{F} -harmonic if $\partial_k c = \partial_{k+1}^* c = 0$. Equivalently, a chain is \mathbb{F} -harmonic if it lies in $Z_k \cap Z^k$. When the underlying field is clear, we may simply say c is harmonic.

Oftentimes, authors working over fields of characteristic 0 will define a chain to be harmonic if it falls within the kernel of the Laplacian $\mathcal{L}_k := \partial_{k+1}\partial_{k+1}^* + \partial_k^*\partial_k$. Over such a field, this definition is equivalent to Definition 2.3, though the two may diverge over a field of positive characteristic. In general, ker $\mathcal{L}_k \supseteq Z_k \cap Z^k$ (see Example 3.2).

CW Complexes. Let X denote a CW complex. We denote the k-skeleton of X by $X^{(k)}$ and the set of k-cells is denoted X_k . Recall that a CW complex is given by iteratively attaching k-cells e_{α}^k as α varies, where each e_{α}^k is homeomorphic to an open Euclidean k-disk. Inducting on the degree k, the k-skeleton is formed from the (k-1)-skeleton via the attaching maps for each k-cell. The attaching map of a cell e_{α}^k is of the form

$$\varphi_{\alpha} \colon S^{k-1}_{\alpha} \to X^{(k-1)},$$

where $\delta e_{\alpha}^{k} = S_{\alpha}^{k-1}$ denotes the boundary of e_{α}^{k} . Additionally, we call a CW complex *degree-wise finite* if it has finitely many cells of each dimension.

All of the topological spaces we are interested in, e.g., simplicial and cubical complexes, admit the structure of a CW complex, including topological graphs.

Cellular Homology. Given a CW complex X and a ring R, there is a chain complex $C_{\bullet}(X; R)$ called the *cellular chain complex*. Each chain module $C_k(X; R)$ is a free R-module with generating set given by the set of k-cells of X, so a typical k-chain is given by an R-linear combination of k-cells. The differentials are determined by the attaching maps of the cell complex via

$$\partial_{k+1}(\alpha) = \sum_{a \in X_k} \langle \delta \alpha, a \rangle a \,,$$

for each (k + 1)-cell α , where $\langle -, - \rangle$ denotes incidence number.

In the case that R is a field and X is degree-wise finite, each cellular chain space $C_k(X; R)$ is finite dimensional with a canonical basis given by the k-dimensional cells of X. This case is our motivating example.

Main Results. We state the first of our main theorems in the general setting of chain complexes of finite dimensional vector spaces over an arbitrary field with chosen bases, though our motivating example is that of a cellular chain complex $C_{\bullet}(X; \mathbb{F})$ of a degree-wise finite CW complex X.

Theorem A. Let C_{\bullet} be a chain complex of finite dimensional vector spaces over a field \mathbb{F} , where each C_k is equipped with a chosen basis and the induced symmetric bilinear form. The following are equivalent.

- (1) (Existence and uniqueness of harmonic representatives) Every homology class in H_k has a unique harmonic representative.
- (2) (Uniqueness of harmonic representatives) For every homology class in H_k with a harmonic representative, that representative is unique.
- (3) The only harmonic representative of the trivial homology class $[0] \in H_k$ is 0.
- (4) $Z_k \cap Z^k \cong H_k$ via the quotient map $z \mapsto [z]$.
- (5) $B_k \cap Z^k = 0.$
- (6) $B_k \oplus Z^k = C_k$.
- (7) $\ker(\partial_k^*\partial_k) = \ker(\partial_k).$
- (8) $\operatorname{im}(\partial_{k+1}^* \partial_{k+1}) = \operatorname{im}(\partial_{k+1}^*).$
- (9) The Moore-Penrose pseudoinverse π^{\dagger} of the projection map $\pi: C_k \to C_k/B_k$ exists.

Furthermore, if these statements hold, then the pseudoinverse π^{\dagger} of (9) takes every homology class to its unique harmonic representative.

Theorem A provides several possible answers to (Q2).

We highlight some notable aspects of Theorem A. Statements (1) through (3) regard existence and uniqueness of harmonic representatives. While uniqueness of harmonic representatives does imply existence, the converse does not hold, as Example 3.2 will demonstrate. Statements (5) and (6) relate to the linear algebraic properties of the chain complex. Note that $B_k^{\perp} = \ker(\partial_k) = \operatorname{im}(\partial_k^*) = Z^k$, so B_k and Z^k are orthogonal complements. If the underlying field has characteristic 0, these spaces are indeed linear complements and statement (5) always holds. Hence the statements of Theorem A hold over any characteristic 0 field. Finally, statement (9) provides an explicit method for computing the unique harmonic representative of a homology class. This is our primary method of computing representatives. These computations, as well as pseudoinverses over arbitrary fields, are discussed in Section 4.

One advantage of stating Theorem A at the level of chain complexes is that it can be applied just as well to a cochain complex, such as the dual complex C^{\bullet} , so long as we take care with the indexing. Recall that a cochain complex is roughly the same as a chain complex, except that the indexing increases as we apply the boundary map instead of decreasing. Since the indexing is completely incidental to the above statements, we can apply Theorem A to C^{\bullet} and get a "dual" theorem. For example, when Theorem A is applied to a cochain complex, statement (5) becomes " $B^k \cap Z_k = 0$," and statements (1) through (4) become statements about harmonic representatives of *co*homology classes. When the statements of Theorem A apply to both a chain complex C_{\bullet} and its dual complex C^{\bullet} , these statements will give us that the harmonic forms are exactly ker \mathcal{L} , as in the classical case.

Definition 2.4. If the statements of Theorem A hold for a cochain complex C_{\bullet} at degree k, we say that C_{\bullet} is homologically harmonic in degree k. If the statements of Theorem A hold for a cochain complex C^{\bullet} at degree k, we say that C^{\bullet} is cohomologically harmonic. We say a degree-wise finite CW complex X is homologically harmonic in degree k over \mathbb{F} when the cellular chain complex $C_{\bullet}(X;\mathbb{F})$ is homologically harmonic, and we say X is cohomologically harmonic in degree k over \mathbb{F} if the dual complex C^{\bullet} is cohomologically harmonic.

Theorem B (Hodge Decomposition over Arbitrary Fields). Let C_{\bullet} be a chain complex of finitedimensional vector spaces with chosen bases, and let $\mathcal{L}_k : C_k \to C_k$ be the discrete Laplacian $\partial \partial^* + \partial^* \partial$. The vector space C_k decomposes into a direct sum of subspaces

$$C_k = \ker \mathcal{L}_k \oplus B_k \oplus B^k$$

if and only if both C_{\bullet} is homologically harmonic in degree k, and the dual complex C^{\bullet} is cohomologically harmonic in degree k.

In this case, the subspaces in the decomposition are orthogonal with respect to the symmetric bilinear form, and the component ker \mathcal{L}_k is equal to $Z_k \cap Z^k$. Furthermore, ker \mathcal{L}_k is isomorphic to both H_k and H^k through the natural maps sending (co)cycles to their respective (co)homology classes.

While Theorem B provides an answer to (Q1) in the general setting of chain complexes, we are most interested in the case when $C_{\bullet} = C_{\bullet}(X; \mathbb{F})$ arises from some degree-wise finite CW complex X. In this case, Theorem B provides necessary and sufficient conditions for a decomposition

$$C_k(X; \mathbb{F}) = \ker(\mathcal{L}_k) \oplus B_k(X; \mathbb{F}) \oplus B^k(X; \mathbb{F}).$$

In Section 5, we focus on the special case of cellular homology of a degree-wise finite CW complex X over finite fields. In this special case, we can ask for which primes p the complex X is homologically harmonic over \mathbb{F}_p . Our approach is to take the matrix sending each rational cycle to its unique harmonic representative and use it to construct a similar map over a finite field \mathbb{F}_p . Since rational homology cannot detect torsion, this approach requires that $H_k(X)$ have no p-torsion. The feasibility of this construction also depends on the entries of the matrix, which we show fall within a certain subring of \mathbb{Q} determined by the combinatorial structure of X. In the process we prove the following result, which we believe to be of independent interest. For an integer N, let $\mathbb{Z}[N^{-1}]$ denote the localization $S^{-1}\mathbb{Z}$, where $S = \{1, N, N^2, N^3, \ldots\}$.



FIGURE 1. A CW complex that fails both existence and uniqueness of \mathbb{F}_2 -harmonic representatives.

Theorem C. For a finite CW complex X, there exists an integer $\Upsilon = \Upsilon(X, k)$ depending on the combinatorial structure of X so that, for any integral homology class in degree k, its rational harmonic representative has coefficients in $\mathbb{Z}[\Upsilon^{-1}]$.

This result is justified by Theorem 5.8, which is stated after we give a rigorous definition of Υ in Definition 5.4. The primary result of Sections 5 and 6 is the following sufficient condition for the homological harmonicity of X.

Theorem D. If $p \nmid \Upsilon(X, k)$, and $H_k(X)$ has no p-torsion, then X is homologically harmonic over \mathbb{F}_p in degree k. That is, every homology class in $H_k(X; \mathbb{F}_p)$ has a unique harmonic representative.

Furthermore, if k = 1 and X is a CW structure on a connected orientable surface, Υ is given by

 $\Upsilon = \det \widehat{\mathcal{L}}_1 / |X_2| \,,$

where $\widehat{\mathcal{L}}_1 = \mathcal{L}_1|_{B_1(X)} : B_1(X; \mathbb{F}_p) \xrightarrow{\cong} B_1(X; \mathbb{F}_p)$ is the Laplacian restricted to the 1-boundaries.

Since only finitely many primes divide Υ , and $H_k(X)$ has nontrivial *p*-torsion for only finitely many primes *p*, we get the following corollary.

Corollary 2.5. Given a degree-wise finite CW complex X and $k \ge 0$, X is homologically harmonic in degree k over \mathbb{F}_p for all but finitely many primes p.

Finally, in Section 7, we put the above theorems to work by computing and plotting some examples of harmonic representatives over various fields. Our experiments will include examples on orientable surfaces and Vietoris Rips complexes.

3. EXISTENCE AND UNIQUENESS OF HARMONIC REPRESENTATIVES

In this section, we explore the properties of a chain complex that affect the existence and uniqueness of harmonic representatives. For a fixed CW complex, a given homology class may have multiple harmonic representatives or none at all depending on the characteristic of the coefficient field. More generally, a homology class of a chain complex may fail to exhibit a harmonic representative, or have many. We begin by considering examples that demonstrate failures of existence and uniqueness.

Example 3.1. Consider a CW complex X with a single 0-cell, two 1-cells denoted a and b, and a single 2-cell denoted E with an attaching map that travels once around a, then once around b. The resulting CW complex, depicted in Figure 1, can be thought of as a cylinder in which a longitudinal line segment has been collapsed. We will consider the homology of this complex with \mathbb{F}_2 coefficients.

Since X is homotopy equivalent to a circle, it has exactly one nontrivial homology class in degree 1, namely the class $[a] = \{a, b\}$. However, neither a nor b is a cocycle, since $\partial_2^*(a) = \partial_2^*(b) = E$. Therefore the homology class [a] has no \mathbb{F}_2 -harmonic representative.

In addition, both cycles in the trivial homology class $[0] = \{0, a + b\}$ are cocycles since $\partial_2^*(a + b) = 2E = 0$, so [0] has two \mathbb{F}_2 -harmonic representatives.



FIGURE 2. A CW complex which fails uniqueness but not existence; every homology class has a non-unique \mathbb{F}_2 -harmonic representative.

Example 3.2. Take the CW complex X from Example 3.1 and subdivide the 2-cell E with an edge e into a pair of 2-cells G and H as pictured in Figure 2; call this new complex X'. We will again consider the homology of X' with \mathbb{F}_2 coefficients. As before, X' has one nontrivial homology class in degree one. There are four cycles in the homology class of b + b', and computing the coboundary of each representative reveals that

$$\partial_2^*(a+a') = \partial_2^*(b+b') = G + H$$
, and
 $\partial_2^*(a+b'+e) = \partial_2^*(a'+b+e) = 0.$

Hence the class of b + b' has two \mathbb{F}_2 -harmonic representatives, namely a + b' + e and a' + b + e.

We can perform a similar computation beginning with the trivial cycle 0. In this case, we find that the class of 0 also has two \mathbb{F}_2 -harmonic representatives, given by 0 and a + a' + b + b'. Additionally, the 1-chain *e* lies in the kernel of the laplacian \mathcal{L}_1 , since $\partial_2 \partial_2^* e = \partial_1^* \partial_1 e = a + b + a' + b'$. This further demonstrates that in general elements of ker \mathcal{L}_k need not be cycles, and that the containment $Z_k \cap Z^k \subseteq \mathcal{L}_k$ can be strict.

3.1. Uniqueness. Given an \mathbb{F} -harmonic representative for a homology class, we can measure its failure to be unique by characterizing the set of all \mathbb{F} -harmonic representatives for its homology class. Suppresing the field \mathbb{F} from the notation, let $\operatorname{Har}(z)$ be the set of all \mathbb{F} -harmonic representatives of $z \in H_k$.

Proposition 3.3. When $\operatorname{Har}(z)$ is non-empty, it forms an additive torsor over $B_k \cap Z^k$. Equivalently, if $\operatorname{Har}(z) \neq \emptyset$, then $\operatorname{Har}(z) = h + (B_k \cap Z^k)$ for any $h \in \operatorname{Har}(z)$. In particular, the \mathbb{F} -harmonic representatives of the trivial homology class coincide with $B_k \cap Z^k$.

Proof. Suppose $h, h' \in \text{Har}(z)$. By definition, $h, h' \in Z^k$ and thus, $h' - h \in Z^k$. Since [h] = [h'] = z, h and h' must differ by a boundary. Hence $h' - h \in B_k \cap Z^k$, showing $h' \in h + (B_k \cap Z^k)$. For the reverse inclusion, let $h \in \text{Har}(z)$ and pick an arbitrary element $h + b \in h + (B_k \cap Z^k)$. Since both h and b are cocycles, so is their sum h + b. Also, [h + b] = [h] = z, so we conclude h + b is a harmonic representative of z. The second statement follows immediately.

In general, $B_k = (Z^k)^{\perp}$ with respect to any symmetric non-degenerate bilinear form on C_k , since $B_k = \operatorname{im}(\partial_{k+1}) = \operatorname{ker}(\partial_{k+1}^*)^{\perp} = (Z^k)^{\perp}$. This means that $B_k \cap Z^k = 0$ if and only if B_k (equivalently, Z^k) is a nondegenerate subspace of C_k . Combined with Proposition 3.3, this gives the following.

Corollary 3.4. Every harmonic representative in C_k is unique if and only if B_k (equivalently Z^k) is a non-degenerate subspace of C_k with respect to the bilinear form on C_k .

Remark 3.5. If \mathbb{F} is a field of characteristic 0, every subspace of C_k is non-degenerate. Corollary 3.4 then implies that harmonic representatives are always unique over a field of characteristic 0.

3.2. Existence. We turn our attention upon the extent to which harmonic representatives exist.

Theorem 3.6. Let Q_k denote the subspace of all degree k-homology classes that have at least one harmonic representative. Then

$$Q_k = \frac{Z_k \cap (Z^k + B_k)}{B_k}$$

The quotient H_k/Q_k is isomorphic to $\frac{Z^k+Z_k}{Z^k+B_k}$, and in particular, the codimension of Q_k in H_k is $\dim(Z^k+Z_k) - \dim(Z^k+B_k)$.

Proof. First we show that $Q_k = \frac{Z_k \cap (Z^k + B_k)}{B_k}$ by dual containment. Let $z + B_k \in Q_k$. Write $z + B_k = h + B_k$ for some harmonic chain h. Since $h \in Z_k \cap Z^k \subseteq Z_k \cap (Z^k + B_k)$, we can conclude $h + B_k \in \frac{Z_k \cap (Z^k + B_k)}{B_k}$. For the reverse inclusion, let $z + B_k \in \frac{Z_k \cap (Z^k + B_k)}{B_k}$, so that $z \in Z_k \cap (Z^k + B_k)$. Since $z \in Z_k$, $z + B_k \in H_k$. Furthermore, z = z' + b for some $z' \in Z^k$, $b \in B_k$. Since both b and z are cycles, z' is a cycle as well. Hence z' is a representative of the homology class of z that is both a cycle and a cocycle. This implies $z + B_k \in Q_k$.

Now that we have shown $Q_k = \frac{Z_k \cap (Z^k + B_k)}{B_k}$, we have

$$H_k/Q_k = \frac{\frac{Z_k}{B_k}}{\frac{Z_k \cap (Z^k + B_k)}{B_k}} \cong \frac{Z_k}{Z_k \cap (Z^k + B_k)}$$

By the second isomorphism theorem,

$$\frac{Z_k}{Z_k \cap (Z^k + B_k)} \cong \frac{Z^k + B_k + Z_k}{Z^k + B_k} = \frac{Z^k + Z_k}{Z^k + B_k}.$$

Remark 3.7. As discussed in Remark 3.5, over a field of characteristic 0, B_k and Z^k are orthogonal complements in C_k . Therefore $Z^k + B_k = C_k$, and $\frac{Z^k + Z_k}{Z^k + B_k} = 0$. Theorem 3.6 then implies that every homology class has a harmonic representative. In addition, by Remark 3.5, that harmonic representative is unique.

4. Computation of Harmonic Representatives

In this section, we focus on computing harmonic representatives by introducing the Moore-Penrose pseudoinverse. We then use the results of Sections 3 and 4 to prove Theorem A, as well as Theorem B, our generalization of the discrete Hodge decomposition.

4.1. Pseudoinverses.

Definition 4.1. Given an $m \times n$ matrix A with entries in \mathbb{F} , the (Moore-Penrose) pseudoinverse of A, denoted A^{\dagger} , is an $n \times m$ matrix satisfying the following:

- (1) $AA^{\dagger}A = A^{\dagger}$.
- (2) $A^{\dagger}AA^{\dagger} = A^{\dagger}.$
- (3) AA^{\dagger} is Hermitian.
- (4) $A^{\dagger}A$ is Hermitian.

The (Moore-Penrose) pseudoinverse generalizes the usual matrix inverse to matrices that fail to be invertible. A preferred solution to the system Ax = y is given by the pseudoinverse: $A^{\dagger}y$. If A is invertible, then $A^{\dagger} = A^{-1}$ and the preferred solution is the unique solution. If Ax = y has multiple solutions, $A^{\dagger}y$ is the solution orthogonal to ker(A), which is also the solution of minimal norm if Ais over a field of characteristic 0. If Ax = y has no solution, then $A^{\dagger}y$ is a solution to the system $Ax = \hat{y}$ where \hat{y} is the orthogonal projection of y onto im(A). Over a field of characteristic 0, this is equivalent to saying that A^{\dagger} minimizes the norm of $y - A(A^{\dagger}y)$.

The following existence and uniqueness theorems are due to Pearl [21, Theorem 1] and Penrose¹ [22, Theorem 1].

 $^{^{1}}$ While Penrose was considering matrices with complex entries, his uniqueness proof works over any field.

Theorem 4.2. Let A be a matrix with entries in a field.

- (1) (Pearl) A pseudoinverse of A exists if and only if $rank(AA^*) = rank(A) = rank(A^*A)$.
- (2) (Penrose) If a pseudoinverse of A does exist, then it is unique.

The uniqueness result justifies the use of the notation A^{\dagger} to denote the pseudoinverse of A.

Remark 4.3. If A is a matrix over a field of characteristic 0, then $\operatorname{rank}(AA^*) = \operatorname{rank}(A^*A) = \operatorname{rank}(A)$ always holds, so A^{\dagger} always exists.

Remark 4.4. If A is full rank, then A^*A is automatically invertible, so by Theorem 4.2, A^{\dagger} exists if and only if $(AA^*)^{-1}$ exists. If this is the case, then it can be shown that $A^{\dagger} = A^*(AA^*)^{-1}$ by directly verifying the pseudoinverse axioms.

Given a linear map $\phi : V \to W$ between finite dimensional vector spaces, the pseudoinverse is traditionally only defined with respect to chosen bases for V and W. In the special case that ϕ is surjective, however, we need only specify a preferred basis for V in order for a pseudoinverse of ϕ to be well-defined.

Definition 4.5. Let $\phi : \mathbb{F}^n \to W$ be a linear map and A a matrix representing ϕ with respect to a basis $B \subseteq W$. If A^{\dagger} exists, then the pseudoinverse of ϕ is defined by $\phi^{\dagger}(w) = A^{\dagger}w$.

We show that this definition does not depend on the choice of basis B. Let B' be another ordered basis for W. Define $\alpha_B : \mathbb{F}^m \to W$ to be the linear isomorphism sending the standard basis to B, and define $\alpha_{B'}$ similarly. Let A and A' be matrices representing ϕ with respect to B and B' respectively, so that

$$\begin{array}{c} \mathbb{F}^n & \xrightarrow{A} & \mathbb{F}^m \\ A' \downarrow & \swarrow & \downarrow \alpha_B \\ \mathbb{F}^m & \xrightarrow{\alpha_{B'}} & W \end{array}$$

commutes. We will show that

(3)

commutes as well. The two paths from W to \mathbb{F}^n in Diagram (3) represent the two possible definitions of ϕ^{\dagger} , each with respect to a different basis for W, so showing Diagram (3) commutes will suffice. By Remark 4.4, $A^{\dagger} = A^* (AA^*)^{-1}$, so

 $\begin{array}{cccc}
\mathbb{F}^n & \stackrel{A^{\dagger}}{\leftarrow} & \mathbb{F}^m \\
A'^{\dagger} & & \uparrow \alpha_B^{-1} \\
\mathbb{F}^m & \stackrel{\bullet}{\leftarrow} & m \\
\end{array}$

$$A^{\dagger} \alpha_B^{-1} = A^* (AA^*)^{-1} \alpha_B^{-1}.$$

By the first diagram, $A = \alpha_B^{-1} \alpha_{B'} A'$. Substituting this into the above expression and simplifying yields $A'^{\dagger} \alpha_{B'}^{-1}$, showing that the diagram commutes.

The following result follows immediately from Remark 4.4.

Proposition 4.6. If $\phi : \mathbb{F}^n \to W$ is a surjective linear transformation, then ϕ^{\dagger} exists if and only if $\phi\phi^*$ is invertible. In this case, $\phi^{\dagger} = \phi^*(\phi\phi^*)^{-1}$.

4.2. Computing the Harmonic Representative. In the special case that each homology class has a unique harmonic representative, that representative can be computed using a pseudoinverse.

Lemma 4.7. Let $\pi: C_k \to C_k/B_k$ be the canonical projection map. Then $\operatorname{im}(\pi^*) = Z^k$.

Proof. By definition of π , we find $B_k = \operatorname{im}(\partial_{k+1}) = \operatorname{ker}(\pi)$. Taking orthogonal complements,

$$\ker(\partial_{k+1}^*) = \operatorname{im}(\partial_{k+1})^{\perp} = \ker(\pi)^{\perp} = \operatorname{im}(\pi^*).$$

Theorem 4.8. Suppose the canonical projection $\pi : C_k \to C_k/B_k$ has pseudoinverse π^{\dagger} . For any $h \in H_k, \pi^{\dagger}(h)$ is a harmonic representative for h.

Proof. By Proposition 4.6, $\pi^{\dagger} = \pi^*(\pi\pi^*)^{-1}$, and in particular this shows that π^{\dagger} is a right inverse of π . Hence $h = \pi\pi^{\dagger}(h) = [\pi^{\dagger}(h)]$, showing that $\pi^{\dagger}(h)$ is a representative of h.

In addition, since $\pi^{\dagger} = \pi^*(\pi\pi^*)^{-1}$, and $(\pi\pi^*)^{-1}$ is full rank, we can see $\operatorname{im}(\pi^{\dagger}) = \operatorname{im}(\pi^*)$, and by Lemma 4.7, $\operatorname{im}(\pi^*) = Z^k$. Therefore $\pi^{\dagger}(h)$ is indeed a cocycle.

Theorem 4.8 allows for the computation of harmonic representatives when π^{\dagger} exists but, as stated in Theorem 4.2, this need not always be the case.

Lemma 4.9. The projection $\pi: C_k \to C_k/B_k$ has a pseudoinverse π^{\dagger} if and only if every harmonic representative is unique.

Proof. As discussed in Remark 4.4, π^{\dagger} exists if and only if $\pi\pi^{*}$ is invertible, and this is true if and only if $im(\pi^{*}) \cap ker(\pi) = 0$. Lemma 4.7 gives $im(\pi^{*}) = Z^{k}$, and clearly $ker(\pi) = B_{k}$. In other words, π^{\dagger} exists exactly when $B_{k} \cap Z^{k} = 0$. Corollary 3.4 finishes the proof.

4.3. Proofs of Theorem A and Theorem B.

Proof of Theorem A. First note that, since $Z_k \cap Z^k$ is the space of harmonic forms, (4) is just an algebraic restatement of (1). Omitting (4) from the list, we prove the equivalence of the remaining items in three parts.

Part 1: $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (5) \Rightarrow (9) \Rightarrow (1)$.

The implications $(1) \Rightarrow (2) \Rightarrow (3)$ are immediate. To see that $(3) \Rightarrow (5)$, recall from Proposition 3.3 that the harmonic representatives of any given homology class form a torsor over $B_k \cap Z^k$, the set of harmonic representatives of the trivial homology class. If 0 is the only such representative, this torsor must be over the trivial space, forcing $B_k \cap Z^k = 0$.

For (5) \Rightarrow (9), π^{\dagger} exists when $\pi\pi^{*}$ is invertible by Proposition 4.6, and since π^{*} is injective, this holds exactly when $\operatorname{im}(\pi^{*}) \cap \operatorname{ker}(\pi) = 0$. Clearly $\operatorname{ker}(\pi) = B_{k}$, and Lemma 4.7 implies $\operatorname{im}(\pi^{*}) = Z^{k}$. Therefore $\operatorname{im}(\pi^{*}) \cap \operatorname{ker}(\pi) = B_{k} \cap Z^{k} = 0$ and so π^{\dagger} exists.

Finally, $(9) \Rightarrow (1)$ since, by Theorem 4.8 and Lemma 4.9, $\pi^{\dagger}(h)$ is the unique harmonic representative of $h \in H_k(X; \mathbb{F})$.

Part 2: (5) and (6) are equivalent.

This follows from a dimension counting argument:

$$\dim(Z^k) = \operatorname{null}(\partial_{k+1}^*) = \dim(C_k) - \operatorname{rank}(\partial_{k+1}^*) = \dim(C_k) - \operatorname{rank}(\partial_{k+1}),$$

where $\operatorname{null}(\partial_{k+1}^*) = \operatorname{dim}(\operatorname{ker}(\partial_{k+1}^*))$. Thus $\operatorname{dim}(B_k) + \operatorname{dim}(Z^k) = \operatorname{dim}(C_k)$.

Part 3: (5), (7), and (8) are equivalent.

Equivalence of (5) and (7) follows from the general fact that, for linear maps f and g, ker $(f \circ g) =$ ker(g) if and only if ker $(f) \cap im(g) = 0$. Choosing $f = \partial_{k+1}^*$ and $g = \partial_{k+1}$ gives the result.

Lastly, items (7) and (8) are related by taking transposes. Indeed,

$$\ker(\partial_{k+1}^*\partial_{k+1}) = \ker(\partial_{k+1}) \iff \ker((\partial_{k+1}^*\partial_{k+1})^{**}) = \ker(\partial_{k+1}^*)$$
$$\iff \operatorname{im}((\partial_{k+1}^*\partial_{k+1})^*)^{\perp} = \operatorname{im}(\partial_{k+1}^*)^{\perp}$$
$$\iff \operatorname{im}(\partial_{k+1}^*\partial_{k+1}) = \operatorname{im}(\partial_{k+1}^*).$$

This completes the proof of equivalence.

The final statement about π^{\dagger} follows from Theorem 4.8, using statement (9) for the existence of π^{\dagger} and statement (2) for uniqueness of the representative.

We finish the section with a proof of the Hodge decomposition.

Proof of Theorem B. Suppose C_{\bullet} is both homologically and cohomologically harmonic in degree k. Applying Theorem A item (6), we decompose C_k as an inner direct sum $C_k = Z^k \oplus B_k$. These subspaces are orthogonal with respect to the standard symmetric bilinear form, since $B_k^{\perp} = Z^k$. Then applying the same part of Theorem A to C^{\bullet} at degree k, we get $C_k = C^k = Z_k \oplus B^k$, which is again an orthogonal decomposition. Intersecting with Z^k gives

$$Z^{k} = (Z^{k} \cap Z_{k}) \oplus (Z^{k} \cap B^{k}) = (Z^{k} \cap Z_{k}) \oplus B^{k}$$

Substituting into the first decomposition above, we obtain the orthogonal decomposition

$$C_k = Z^k \oplus B_k = (Z^k \cap Z_k) \oplus B^k \oplus B_k$$

as desired.

Conversely, suppose the above decomposition holds. Then in particular, $(Z_k \cap Z^k) \cap B_k = B_k \cap Z^k = 0$. Theorem A item (5) implies C_{\bullet} is homologically harmonic. A similar argument shows that $B^k \cap Z_k = \emptyset$, implying that C_{\bullet} is cohomologically harmonic as well, completing the left hand implication.

Next we show that $Z_k \cap Z^k = \ker \mathcal{L}_k$. Since $\mathcal{L}_k = \partial_{k+1}\partial_{k+1}^* + \partial_k^*\partial_k$, clearly $Z_k \cap Z^k \subseteq \ker \mathcal{L}_k$. For the opposite inclusion, suppose $\mathcal{L}_k x = 0$. Then $\partial_k^*\partial_k x = \partial_{k+1}\partial_{k+1}^*(-x) \in B_k \cap B^k = 0$. Applying item 7 to both C_{\bullet} and C^{\bullet} shows $x \in Z_k \cap Z^k$ as desired.

We have already shown that such a decomposition is orthogonal, so it only remains to show that the natural maps from $(Z^k \cap Z_k)$ to H_k and H^k are both isomorphisms. This follows from applying Theorem A item (4) to both C_{\bullet} and C^{\bullet} .

5. HARMONIC REPRESENTATIVES ON CW COMPLEXES OVER \mathbb{F}_p

We have explored when a chain complex C_{\bullet} is homologically harmonic in degree k. However, in the case that $C_{\bullet} = C_{\bullet}(X; \mathbb{F})$, it is not clear how changing the field \mathbb{F} affects this property. In this section, we examine the special case where $\mathbb{F} = \mathbb{F}_p$ is a finite field of prime order, examining for which p a fixed CW complex is homologically harmonic over \mathbb{F}_p . Our conditions are based on the integer Υ , which is itself defined in terms of the following combinatorial structure.

Definition 5.1 ([7, Definition 1.10]). Let X be a finite CW complex with dimension at least k. A k-dimensional spanning cotree of X is a subcomplex L such that

- $X^{(k-1)} \subset L \subset X^{(k)}$
- The inclusion $i_L: L \to X$ induces rational isomorphisms

$$i_{L*}: H_k(L; \mathbb{Q}) \to H_k(X; \mathbb{Q}) \text{ and } i_{L*}: H_{k-1}(L; \mathbb{Q}) \to H_{k-1}(X; \mathbb{Q}).$$

Remark 5.2. Recall that the rows of the matrix $\partial_k : C_k(X) \to C_{k-1}(X)$ correspond to the k-cells of X. We could equivalently define a k-dimensional spanning cotree to be a subcomplex $X^{(k-1)} \subseteq L \subseteq X^{(k)}$ such that the k-cells not in L form a row basis for the matrix ∂_k .

Spanning co-trees of the appropriate dimension always exist [7, Lemma 2.1]. Furthermore, the rational isomorphisms in their definition imply the relative integral homology group $H_k(X, L)$ is always finite for any k-dimensional spanning cotree L.

Definition 5.3. Given a k-dimensional spanning cotree L, the weight of L is

$$a_L = |H_k(X, L)|.$$

Definition 5.4. Given a finite CW complex X and $k \ge 0$, define

$$\Upsilon = \Upsilon(X,k) = \left(\sum_{L} a_{L}^{2}\right) \prod_{L} a_{L} ,$$

where the sum and product are each over all k-dimensional spanning cotrees of X.

Definition 5.5. Let $\mathbb{Z}_{(p)} \subseteq \mathbb{Q}$ be the subring of rational numbers with denominators not divisible by p. Define $r_p : \mathbb{Z}_{(p)} \to \mathbb{F}_p$ to be the ring map $r_p(a/b) = ab^{-1} \mod p$. Given a matrix A with entries in $\mathbb{Z}_{(p)}$, define R_pA to be the matrix obtained by applying r_p entrywise to A.

Lemma 5.6. Let p be a prime not dividing any denominators of the entries of $\pi^{\dagger}\pi$, and such that $H_k(X)$ has no p-torsion. Then the matrix $R_p(\pi^{\dagger}\pi)$ is the orthogonal projection map of $C_k(X; \mathbb{F}_p)$ onto $Z^k(X; \mathbb{F}_p)$ with respect to the standard bilinear form b_k on $C_k(X; \mathbb{F}_p)$. That is,

- $R_p(\pi^{\dagger}\pi)$ is self-adjoint under b_k .
- $R_p(\pi^{\dagger}\pi)$ fixes its image.
- $\operatorname{im}(R_p(\pi^{\dagger}\pi)) = Z^k$.

Proof. First, it is clear that since $\pi^{\dagger}\pi$ is a symmetric matrix, so is $R_p(\pi^{\dagger}\pi)$. A symmetric matrix is self-adjoint under the standard bilinear form b_k if and only if it is symmetric, so $R_p(\pi^{\dagger}\pi)$ is indeed self-adjoint.

We show that $\operatorname{im}(R_p(\pi^{\dagger}\pi)) \subseteq Z^k(X; \mathbb{F}_p)$. If we let $c \in C_k(X; \mathbb{F}_p)$ and pick $c' \in C_k(X; \mathbb{Q})$ with $r_{p*}(c') = c$, then we have

$$R_p(\pi^{\dagger}\pi)(c) = r_{p*}(\pi^{\dagger}\pi(c'))$$

Since $\pi^{\dagger}\pi(c')$ is always a cocycle, and r_{p*} is a chain map, we get that $r_{p*}(\pi^{\dagger}\pi(c')) \in Z^k(X; \mathbb{F}_p)$.

Finally, we show $R_p(\pi^{\dagger}\pi)$ fixes every $z \in Z^k(X; \mathbb{F}_p)$. We claim that there is a $z' \in Z^k(X; \mathbb{Z}[\Upsilon^{-1}])$ such that $r_{p*}(z') = z$. To see this, consider the short exact sequence of groups

$$0 \to \mathbb{Z}[\Upsilon^{-1}] \xrightarrow{\cdot p} \mathbb{Z}[\Upsilon^{-1}] \xrightarrow{r_p} \mathbb{F}_p \to 0,$$

and the induced short exact sequence of cochain complexes

$$0 \to C^k(X; \mathbb{Z}[\Upsilon^{-1}]) \xrightarrow{p} C^k(X; \mathbb{Z}[\Upsilon^{-1}]) \xrightarrow{r_{p*}} C^k(X; \mathbb{F}_p) \to 0.$$

The snake lemma gives us the exact sequence

$$Z^{k}(X; \mathbb{Z}[\Upsilon^{-1}]) \xrightarrow{r_{p*}} Z^{k}(X; \mathbb{F}_{p}) \xrightarrow{\delta} C^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}]) / B^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}]).$$

To prove that our desired z' exists, by exactness it suffices to show that $\delta = 0$. Indeed, $Z^k(X; \mathbb{F}_p)$ consists entirely of *p*-torsion, so it further suffices to show that $C^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}])/B^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}])$ has no *p*-torsion. A representative of a *p*-torsion element would be a chain $x \in C^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}])$ with $px = \partial^* y$ for some y. Then px is a cocycle, and hence so is x. We now have that x represents a *p*-torsion class of $H^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}])$. By an application of the universal coefficient theorem, since $p \nmid \Upsilon$, the *p*-torsion of $H^{k+1}(X; \mathbb{Z}[\Upsilon^{-1}])$ is isomorphic to the *p*-torsion of $H_k(X)$, and we have assumed that the only such homology class is trivial. Therefore $\delta = 0$, and we can pick a cocycle z' with $r_{p*}(z') = z$.

$$R_p(\pi^{\dagger}\pi)(z) = r_{p*}(\pi^{\dagger}\pi(z')) = r_{p*}(z') = z,$$

which completes the proof.

Corollary 5.7. If p does not divide the denominators of the entries of the matrix $\pi^{\dagger}\pi$ and $H_k(X)$ has no p-torsion, then $C_k(X; \mathbb{F}_p) = Z^k(X; \mathbb{F}_p) + B_k(X; \mathbb{F}_p)$, and X is homologically harmonic in degree k over \mathbb{F}_p .

Proof. Pick any chain $c \in C_k(X; \mathbb{F}_p)$. Then

$$c = (c - R_p(\pi^{\dagger}\pi)(c)) + R_p(\pi^{\dagger}\pi)(c)$$

By the previous result, $R_p(\pi^{\dagger}\pi)(c)$ is a cocycle. To show that $c - R_p(\pi^{\dagger}\pi)(c)$ is a boundary, we write

$$R_{p}(\pi^{\dagger}\pi)(c - R_{p}(\pi^{\dagger}\pi)(c)) = R_{p}(\pi^{\dagger}\pi)(c) - R_{p}(\pi^{\dagger}\pi) \circ R_{p}(\pi^{\dagger}\pi)(c).$$

Again by the previous result, $R_p(\pi^{\dagger}\pi)$ fixes its image, so the above expression is 0. Then $c - R_p(\pi^{\dagger}\pi)(c)$ is a boundary, since it is orthogonal to every cocycle z:

$$b_k(z, c - R_p(\pi^{\dagger}\pi)(c)) = b_k(z, c) - b_k(z, R_p(\pi^{\dagger}\pi)(c))$$

= $b_k(z, c) - b_k(R_p(\pi^{\dagger}\pi)(z), c)$
= $b_k(z, c) - b_k(z, c) = 0.$

Hence $C_k(X; \mathbb{F}_p) = Z^k(X; \mathbb{F}_p) + B_k(X; \mathbb{F}_p)$. By Theorem A, it follows that X is homologically harmonic in degree k over \mathbb{F}_{p} .

The conditions of Corollary 5.7 depend on the denominators of $\pi^{\dagger}\pi$. The following proposition gives a combinatorial description of the denominators that can appear.

Theorem 5.8. The entries of the matrix $\pi^{\dagger}\pi$ lie in $\mathbb{Z}[\Upsilon^{-1}]$.

Proof. By clearing denominators (or choosing an appropriate basis for $C_k(X;\mathbb{Q})/B_k(X;\mathbb{Q})$), we may assume the entries of π are integers. Following [7, Thm A] as adapted from the general construction of [5, Thm 2.1], we can give an explicit construction of π^{\dagger} . Let *m* be the rank of $C_k(X;\mathbb{Q})/B_k(X;\mathbb{Q})$ and for a subset $S \subset \{1, 2, \ldots, |X_k|\}$ of cardinality m, let π_S denote the $m \times m$ submatrix of π whose rows correspond to indices in S. We restrict our focus to only those S so that π_S is invertible. Denote the inclusion corresponding to those rows by $i_S: \mathbb{Z}^m \to C_k(X)$, set $t_S = \det(\pi_S)^2$, and finally set $\nabla = \sum_{S} t_{S}$. Then

$$\pi^{\dagger} = \frac{1}{\nabla} \sum_{S} t_{S} i_{S} (\pi_{S})^{-1} \,.$$

Under the correspondence described in [7, Thm A, Rem 2.7], each such S corresponds to a spanning cotree L and furthermore, $t_S = a_L^2$. The inverse $i_S(\pi_S)^{-1}$ requires the invertibility of a_L and thus by definition of Υ , π^{\dagger} has coefficients in $\mathbb{Z}[\Upsilon^{-1}]$. Hence the product $\pi^{\dagger}\pi$ has coefficients in $\mathbb{Z}[\Upsilon^{-1}]$.

6. HARMONIC REPRESENTATIVES ON ORIENTABLE SURFACES

A given CW complex X may have many spanning cotrees of a given dimension, since the corresponding boundary map may have many row bases. This can make Υ difficult to compute directly. In the case that X = M is a CW structure on an orientable surface, the formula for Υ can be simplified, making its computation simpler. As we will show, for any spanning cotree L of M, the relative homology group $H_1(M,L)$ is trivial, in which case the formula for Υ simplifies to just the number of spanning cotrees. We take the convention that a surface is connected.

6.1. The Dual Cell Structure. Given a CW decomposition M of an orientable surface, there are multiple ways of constructing the dual complex M^* [13]. A careful construction of the dual cell structure is beyond the scope of this paper. The salient points we need are the following.

- There is a canonical bijection between the k-cells of M and the (2-k)-cells of M^* for $0 \le k \le 2$.
- This lets us identify $C_k(M)$ and $C_{2-k}(M^*)$. Under the above identifications, $\partial_1^{M^*}: C_1(M^*) \to C_0(M^*)$ is the transpose of $\partial_2^M: C_2(M) \to C_1(M)$, and similarly $\partial_2^{M^*}: C_2(M^*) \to C_1(M^*)$ is the transpose of $\partial_1^M: C_1(M) \to C_0(M)$.

6.2. Connections between Spanning Trees and Cotrees. In Remark 5.2 we note that a spanning cotree can be defined as corresponding to the complement of a row basis for the appropriate boundary matrix. We can similarly define a k-dimensional spanning tree T to be a subcomplex $X^{(k-1)} \subseteq T \subseteq$ $X^{(k)}$ such that the k-cells of T correspond to a column basis for ∂_k (see [6] for a homological definition). When X is a connected graph and k = 1, this definition aligns with the usual graph-theoretic notion of a spanning tree. When it is understood that T is a k-dimensional spanning tree, we often identify

T with its set of k-cells $T_k \subset X_k$. Given a subcomplex $X^{(k-1)} \subseteq L \subseteq X^{(k)}$ such as a spanning tree or cotree, we use L^{\perp} to denote the subcomplex $X^{(k-1)} \subseteq L^{\perp} \subseteq X^{(k)}$ whose k-cells are $X_k \setminus L_k$. If we identify L with its set of k-cells, then L^{\perp} is simply the complement of L. In the special case that L is a 1-dimensional spanning correction of L. of a CW structure M on an orientable manifold, L^{\perp} can be thought of as a 1-dimensional spanning tree of the dual cell structure M^* . This is because if L is a spanning cotree, then its 1-cells correspond to the complement of a row basis for ∂_2^M . Hence the 1-cells of L^{\perp} correspond to a column basis of $(\partial_2^M)^* = \partial_1^{M^*}.$

Definition 6.1. Given a k-dimensional spanning tree T, its weight is defined to be

$$\theta_T = |H_{k-1}(T)_t|.$$

The following result can simplify calculations of weights of spanning trees and cotrees. Let $X_{k,k-2} := X^{(k)}/X^{(k-2)}$, the k-skeleton of X with the (k-2)-skeleton collapsed to a point.

Proposition 6.2 ([7, Lemma 4.2]). For any spanning tree T or spanning cotree L,

$$\theta_T = \theta_{T_{k,k-2}} \qquad a_L = a_{L_{k,k-2}}.$$

The following formula from the same paper will help us relate these weights to the restricted Laplacian.

Proposition 6.3 ([7, Theorem C] Higher Matrix-Tree Theorem). Let $\widehat{\mathcal{L}}_k : B_k(X) \to B_k(X)$ be the Laplacian restricted to the k-boundaries. Then

$$\det \widehat{\mathcal{L}}_k = \frac{1}{\theta_X^2} \left(\sum_L a_L^2 \right) \left(\sum_T \theta_T^2 \right) \,,$$

where the first sum ranges over all k-dimensional spanning cotrees of X, and the second ranges over (k+1)-dimensional spanning trees.

6.3. Smith Normal Form. Given an $m \times n$ integer matrix A, its Smith normal form is another $m \times n$ matrix SAT where $S \in GL_m(\mathbb{Z}), T \in GL_n(\mathbb{Z})$, and SAT is of the form

a_1	0	0		$0 \rangle$
0	a_2	0		0
0	0	a_3		0
.				
	:	:	••	
$\int 0$	0	0		0/

with $a_i | a_{i+1}$ for all *i*. The Smith normal form of *A* always exists and is unique up to the signs of the nonzero entries. The crucial properties we will be using are that the Smith normal forms of *A* and A^T are themselves transposes, and that $|\operatorname{coker}(A)| = |\operatorname{coker}(SAT)| = |a_1 \dots a_k|$.

Lemma 6.4. Let T be a k-dimensional spanning tree of X, and let t_1, \ldots, t_n be the nonzero entries of the Smith normal form of the composition $C_k(T) \hookrightarrow C_k(X) \xrightarrow{\partial} C_{k-1}(X)$. Then

$$\theta_T = t_1 \cdots t_n.$$

Similarly, if L is a (k-1)-dimensional spanning cotree of X, and ℓ_1, \ldots, ℓ_n are the nonzero entries of the Smith normal form of $C_k(X) \xrightarrow{\partial} C_{k-1}(X) \twoheadrightarrow C_{k-1}(X)/C_{k-1}(L)$, then

$$a_L = \ell_1 \cdots \ell_n.$$

Remark 6.5. The matrix for $C_k(T) \hookrightarrow C_k(X) \xrightarrow{\partial} C_{k-1}(X)$ is the boundary matrix for X restricted to the columns corresponding to the k-cells of T, precisely the submatrix mentioned in our definition of a spanning tree. Similarly, the matrix for $C_k(X) \xrightarrow{\partial} C_{k-1}(X) \twoheadrightarrow C_{k-1}(X)/C_{k-1}(L)$ is the boundary matrix for X after removing the rows corresponding to the (k-1)-cells of L, which is the submatrix mentioned by Remark 5.2.

Proof of 6.4. We handle the spanning tree first. We know that $\theta_T = \theta_{T_{k,k-2}} = |H_{k-1}(T)_t|$ by Proposition 6.2, which is the size of the torsion part of the (k-1)-homology of the chain complex

$$\cdots \to 0 \to C_k(T) \xrightarrow{\partial_k^T} C_{k-1}(T) \to 0 \to \dots,$$

whose homology is $\operatorname{coker}(\partial_k^T)$. Note that $C_{k-1}(T) = C_{k-1}(X)$, and the map ∂_k^T is the boundary map $\partial : C_k(X) \to C_{k-1}(X)$ restricted to $C_k(T)$. Therefore we can factor the map as the composition in the statement of the lemma. Our remark about the size of the cokernel in Subsection 6.3 finishes this part.

The spanning cotree is handled similarly. Using the fact that $a_L = a_{L_{k,k-2}}$, it suffices to calculate the k-1 homology of the relative chain complex

$$\cdots \to 0 \to C_k(X)/C_k(L) \xrightarrow{\partial_k} C_{k-1}(X)/C_{k-1}(L) \to 0 \to \dots$$

which is $\operatorname{coker}(\overline{\partial}_k)$. We already know this cokernel is finite, since its order is a_L , so its free part must be trivial. Note that $C_k(L) = 0$, so the domain of this map is $C_k(X)$. It then factors as the composition in the statement of the lemma. Since the entire cokernel is torsion, we can again use our remark about the size of the cokernel in Subsection 6.3 to finish the proof.

For the statement of the following lemma, recall that if L is a 1-dimensional spanning cotree of M, we can regard L^{\perp} as a 1-dimensional spanning tree of M^* .

Lemma 6.6. If L is a spanning cotree for M of any dimension $0 \le k \le 2$, then $a_L = \theta_{L^{\perp}}$.

Proof. The matrices for $C_{k+1}(M) \xrightarrow{\partial_{k+1}^M} C_k(M) \twoheadrightarrow C_k(M)/C_k(L)$ and $C_k(L^{\perp}) \hookrightarrow C_k(M^*) \xrightarrow{\partial_k^M^*} L^{-1}(X)$ are transposes of each other. Indeed, the former is the matrix for ∂_{k+1}^M with the rows corresponding to the 1-cells of L removed, while the latter is the matrix for $\partial_k^{M^*}$ (the transpose of ∂_{k+1}^M) with the columns corresponding to the 1-cells of L removed. Since these matrices are transposes, we conclude that their Smith normal forms are transposes as well, and in particular have the same nonzero entries. Lemma 6.4 finishes the proof.

Lemma 6.7. For any 1-dimensional spanning cotree L of M, $a_L = 1$.

Proof. By Lemma 6.6, it suffices to prove that every 1-dimensional spanning tree T of M^* has weight 1. This follows since zero-degree homology is always free.

Theorem 6.8. Let M be a CW structure on an orientable surface. Then $\Upsilon(M, 1)$ is the number of 1-dimensional spanning cotrees of M, which equals

$$\Upsilon(M,1) = \det \widehat{\mathcal{L}}_1 / |M_2|$$

where $\widehat{\mathcal{L}}_1: B_1(X) \to B_1(X)$ is the restricted Laplacian, and $|M_2|$ is the number of 2-cells of M.

Proof. The fact that $\Upsilon(M, 1)$ equals the number of spanning cotrees of M follows immediately from Lemma 6.7 and the definition of $\Upsilon(M, 1)$:

$$\Upsilon(M,1) = \left(\sum_{L} |H_1(X,L)|^2\right) \prod_{L} |H_1(X,L)| = \left(\sum_{L} a_L^2\right) \prod_{L} a_L = \sum_{L} 1.$$

Rearranging the equation of Proposition 6.3 gives the expression

$$\Upsilon(M,1) = \frac{\theta_X^2 \det \widehat{\mathcal{L}}_1}{\sum_T \theta_T^2}$$

Since X is an orientable surface, its second homology has no torsion and $\theta_X = 1$. Additionally, we can rewrite each θ_T as $a_{T^{\perp}}$, the weight of a zero-dimensional spanning cotree in M^* . We can derive from the definition of a spanning cotree that any spanning cotree of a connected complex consists of a single vertex. Hence $\theta_T = a_{L^{\perp}} = 1$. Plugging these into the equation above gives

$$\Upsilon(M,1) = \frac{\det \widehat{\mathcal{L}}_1}{|(M^*)_0|} = \frac{\det \widehat{\mathcal{L}}_1}{|M_2|}.$$

Corollary 6.9. Let M be a CW structure on an orientable surface and p a prime not dividing det $\hat{\mathcal{L}}_1/|M_2|$. Then M is homologically harmonic in degree 1 over \mathbb{F}_p .



FIGURE 3. A harmonic representative on the torus. (Top) A minimal triangulation X of the torus, with 7 vertices, 21 edges, and 14 faces. The vertices and edges along the top edge of the square are identified with those on the bottom, and similarly with the left and right edges. (Lower left) The cycle $x \in Z_1(X; \mathbb{F}_2)$. The edges in its support are emphasized in blue, with all other edges drawn dashed. (Lower right) The unique harmonic representative of the homology class [x], drawn in the same fashion.

7. Experiments

In this section we present visualizations of harmonic representatives over various fields. Much of the underlying SageMath code from this section is available as a Python package called simpleHarmony. The source code, installation instructions, and examples can be found on github [26]. Our examples come from two common applications of simplicial complexes: triangulations of surfaces and point cloud datasets. There are several methods of constructing a simplicial complex on the latter, including the Čech complex and alpha complex. We elect to use the Vietoris-Rips complex, which is the clique complex on a proximity graph of the data points. See [12] for a friendly introduction to Čech and Vietoris-Rips complexes, and their applications to persistent homology.

With a simplicial complex X in hand, we manually choose an integral cycle $z \in Z(X;\mathbb{Z})$ which appears to represent a topological feature of interest. We rationalize this cycle by mapping it to $Z(X;\mathbb{Q})$, where we then compute and plot its harmonic representative using the pseudoinverse from Theorem 4.8. For comparison, we also map the integral cycle to $Z(X;\mathbb{F}_p)$ for some prime p and calculate its harmonic representative using the same method. As highlighted by Lemma 4.9, the pseudoinverse need not exist over \mathbb{F}_p for all primes p, so we search for the smallest prime p such that the necessary pseudoinverse exists. After finding such a prime, we proceed to compute and plot the harmonic representative.

7.1. Minimal Triangulation of the Torus. Consider the minimal triangulation of the torus depicted in Figure 3 (top). We examine the first degree homology and harmonic representatives of this simplicial complex X over the field \mathbb{F}_2 . Using code written in sage, we verify that X is homologically harmonic over \mathbb{F}_2 in degree 1. This allows us to pick a particular cycle $x \in Z_1(X; \mathbb{F}_2)$ (Figure 3, bottom left) and compute the unique harmonic representative of [x] (Figure 3, bottom right). We perform this computation by calculating the pseudoinverse from Theorem A statement (9) and applying it to [x].

Alternatively, we can use the methods of Section 5 to find the same representative. Indeed, we use sage to compute the determinant of the restricted Laplacian and, using Theorem D, arrive at

 $\Upsilon(X,1) = 3 \cdot 7^5 = 50421$. Since Υ is odd and p = 2, Theorem D also ensures that X is homologically harmonic in degree k over \mathbb{F}_p . Furthermore, Theorem 5.8 ensures that the entries of the rational matrix $\pi^{\dagger}\pi$ fall within $\mathbb{Z}[\Upsilon^{-1}]$. Applying $R_p(\pi^{\dagger}\pi)$ to x then gives the same \mathbb{F}_2 -harmonic representative as before, showing that the two methods do indeed agree as expected.

7.2. Point Cloud Sampled from a Lemniscate. For this experiment, we randomly sample 50 points from a lemniscate (a self-intersecting smooth curve homeomorphic to the wedge product $S^1 \vee S^1$) embedded within \mathbb{R}^2 with Gaussian noise. We construct the Vietoris-Rips complex X on this dataset with radius 0.45 (Figure 4, top left). We pick an integral cycle of interest $z \in H_1(X;\mathbb{Z})$, by selecting a cycle whose support wraps around the right hand side (Figure 4, top right). We then consider z as a cycle over both \mathbb{Q} and \mathbb{F}_5 , computing its harmonic representative in each case (Figure 4, middle left and right). In the latter case, we achieve this with the pseudoinverse from Theorem A item (9). To highlight the similarity between these harmonic representatives, we also plot the symmetric difference of their supports (Figure 4, bottom).

In this case, the harmonic representatives over both \mathbb{Q} and \mathbb{F}_5 seem to visualize the initial geometric feature, with the support of each lying most heavily on the right hand circle. They both localize the topological phenomenon we are trying to isolate. This highlights one potential application of harmonic representatives over finite fields to topological data analysis, allowing practitioners another way to visualize persistent homology classes.

7.3. Point Cloud Sampled from $S^1 \vee S^2 \vee S^1$. We next run a similar experiment on a larger dataset in \mathbb{R}^3 , inspired by [10, Section 3.1], wherein we sample 130 points from $S^1 \vee S^2 \vee S^1$, a sphere with two circles attached at antipodal points. Of those 130 points, we take 70 from the sphere and 30 from each circle, with Gaussian noise of standard deviation 0.02. We construct the Vietoris-Rips complex X on this dataset with radius 1.3 and choose an integral cycle $z \in H_1(X;\mathbb{Z})$ whose support encompasses the circle attached to the right hand side of the sphere (Figure 5, top left). As in the previous experiment, we plot the harmonic representatives of z when considered as a cycle over \mathbb{Q} and a finite field (Figure 5, top right, bottom left). This time, the smallest field over which our calculation succeeded was \mathbb{F}_{47} , showing that X is not homologically harmonic over any smaller primes.

Notice that the support of the harmonic representative over the finite field is much larger in this case than the previous experiment. Indeed, of the 1948 1-simplices in the Vietoris-Rips complex, 1893 have a nonzero coefficient in this cycle, which is not far from the 1907 that we would expect from a randomly chosen 1-chain. Note, however, that the support of the rational harmonic form is similarly large, as shown by the relatively sparse symmetric difference of the supports (Figure 5, bottom right). These results also differ from the previous experiment in that we require a much larger finite field before the statements of Theorem A hold.

References

- Karim Adiprasito, June Huh, and Eric Katz. Hodge theory for combinatorial geometries. Annals of Mathematics, 188(2):381–452, 2018.
- [2] Matthew Baker. Hodge theory in combinatorics. Bulletin of the American Mathematical Society, 55(1):57-80, 2018.
- [3] Saugata Basu and Nathanael Cox. Harmonic persistent homology, 2021.
- [4] Ulrich Bauer. Ripser: efficient computation of vietoris-rips persistence barcodes. Journal of Applied and Computational Topology, 5(3):391-423, Sep 2021.
- [5] Aharon Ben-Tal and Marc Teboulle. A geometric property of the least squares solution of linear equations. *Linear algebra and its applications*, 139:165–170, 1990.
- [6] Michael J. Catanzaro, Vladimir Y. Chernyak, and John R. Klein. Kirchhoff's theorems in higher dimensions and Reidemeister torsion. *Homology, Homotopy and Applications*, 17(1):165–189, 2015.
- [7] Michael J. Catanzaro, Vladimir Y. Chernyak, and John R. Klein. A higher Boltzmann distribution. Journal of Applied and Computational Topology, 1(2):215–240, 2017.
- [8] Isaac Chavel. Eigenvalues in Riemannian geometry. Academic press, 1984.
- [9] Persi Diaconis. A generalization of spectral analysis with application to ranked data. The Annals of Statistics, pages 949–979, 1989.
- [10] Stefania Ebli and Gard Spreemann. A notion of harmonic clustering in simplicial complexes. 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), Dec 2019.



FIGURE 4. A Vietoris-Rips complex on a point cloud dataset sampled from a lemniscate with noise (upper left), a chosen integral cycle in the Vietoris-Rips complex (upper right), the corresponding harmonic forms over \mathbb{Q} and \mathbb{F}_5 (middle left and right), and the symmetric difference of those two supports.



FIGURE 5. An integral cycle in the Vietoris-Rips complex on a point cloud sampled from $S^1 \vee S^2 \vee S^1$ with noise (upper left), the corresponding harmonic representatives over \mathbb{Q} and \mathbb{F}_{47} (upper right and lower left), and the symmetric difference of those two supports.

- Beno Eckmann. Harmonische funktionen und randwertaufgaben in einem komplex. Commentarii Mathematici Helvetici, 17(1):240–255, Dec 1944.
- [12] Robert Ghrist. Barcodes: The persistent topology of data. Bull. Amer. Math., 45(1):61–75, 2008.
- [13] Allen Hatcher. Algebraic Topology. Cambridge University Press, 2002.
- [14] William V.D. Hodge. The theory and applications of harmonic integrals. Cambridge University Press, 1941.
- [15] Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. Statistical ranking and combinatorial Hodge theory. Mathematical Programming, 127(1):203-244, 2011.
- [16] Mark Kac. Can one hear the shape of a drum? The American Mathematical Monthly, 73(4P2):1–23, 1966.
- [17] Sara Kališnik, Vitaliy Kurlin, and Davorin Lešnik. A higher-dimensional homologically persistent skeleton. Advances in Applied Mathematics, 102:113–142, 2019.
- [18] Lek-Heng Lim. Hodge laplacians on graphs. Siam Review, 62(3):685–715, 2020.
- [19] Sayan Mukherjee and John Steenbergen. Random walks on simplicial complexes and harmonics. Random Structures & Algorithms, 49(2), September 2016.
- [20] Facundo Mémoli, Zhengchao Wan, and Yusu Wang. Persistent laplacians: properties, algorithms and implications, 2021.
- [21] Martin H. Pearl. Generalized inverses of matrices with entries taken from an arbitrary field. Linear Algebra and its Applications, 1(4):571 587, 1968.
- [22] R. Penrose. A generalized inverse for matrices. Mathematical Proceedings of the Cambridge Philosophical Society, 51(3):406–413, 1955.
- [23] Steven Rosenberg. The Laplacian on a Riemannian Manifold: An Introduction to Analysis on Manifolds. Cambridge University Press, January 1997.
- [24] Donald E Taylor. The geometry of the classical groups, volume 9. Heldermann Verlag, 1992.
- [25] Sebastiano Vigna. Spectral ranking. Network Science, 4(4):433-445, 2016. Publisher: Cambridge University Press.
- [26] Brantley Vose. simpleHarmony: A sage package for finding discrete harmonic forms on simplicial complexes.
- [27] Rui Wang, Duc Duy Nguyen, and Guo-Wei Wei. Persistent spectral graph. International journal for numerical methods in biomedical engineering, 36(9):e3376–e3376, Sep 2020. 32515170[pmid].

- [28] Frank W. Warner. Foundations of Differentiable Manifolds and Lie Groups. Graduate Texts in Mathematics. Springer-Verlag, New York, 1983.
- [29] E. L. Wilmer, David A. Levin, and Yuval Peres. Markov chains and mixing times, volume 107. American Mathematical Soc, Providence, R.I, 2009.