Hodge Laplacians on Graphs

Lek-Heng Lim

ABSTRACT. This is an elementary introduction to the Hodge Laplacian on a graph, a *higher-order* generalization of the graph Laplacian. We will discuss basic properties including cohomology and Hodge theory. At the end we will also discuss the nonlinear Laplacian on a graph, a *nonlinear* generalization of the graph Laplacian as its name implies. These generalized Laplacians will be constructed out of coboundary operators, i.e., discrete analogues of exterior derivatives. The main feature of our approach is simplicity — this article requires only knowledge of linear algebra and graph theory.

1. Introduction

The primary goal of this article is to introduce readers to the Hodge Laplacian on a graph and discuss some of its properties, notably the Hodge decomposition. To understand its significance, it is inevitable that we will also have to discuss the basic ideas behind cohomology, but we will do so in a way that is as elementary as possible and with a view towards applications in the information sciences.

If the classical Hodge theory on Riemannian manifolds [**41, 67**] is "differentiable Hodge theory," the Hodge theory on metric spaces [**7, 61**] "continuous Hodge theory," and the Hodge theory on simplicial complexes [**29, 31**] "discrete Hodge theory," then the version here may be considered "graph-theoretic Hodge theory."

Unlike physical problems arising from areas such as continuum mechanics or electromagnetics, where the differentiable Hodge–de Rham theory has been applied with great efficacy for both modeling and computations [**3, 30, 45, 53, 65, 66, 70**], those arising from data analytic applications are likely to be far less structured [**4, 15, 16, 26, 40, 43, 56, 69**]. Often one could at best assume some weak notion of proximity of data points. The Hodge theory introduced in this article requires nothing more than the data set having the structure of an undirected graph and is conceivably more suitable for non-physical applications such as those arising from the biological or information sciences. We will discuss two examples of such applications, to ranking and to game theory.

Readership. To some mathematical cognoscenti, cohomology is like calculus — an everyday tool used without a second thought. This article is not for them. It is intended for those who, like the author, have attempted to learn about cohomology from standard sources in algebraic topology but found them impenetrable and gave up. However, the author has had the good fortune to interact with some of those cognoscenti and did ultimately pick up bits and pieces through diffusion. One goal of this article is to share with a wider audience these bits and pieces that he believes make up the gist of the idea, which is remarkably simple.

This article also serves as an expanded written record of the author's lecture in the AMS Short Course on *Geometry and Topology in Statistical Inference* at the 2014 Joint Mathematics Meetings. The target audience, as per AMS's guidelines [**1**], comprises:

- the idly curious, knowing little or nothing specific of the field beyond a layman's or graduate student's familiarity;
- peripheralists, who have read a few articles, perhaps dabbled once or twice in the field, and would like to have a perspective of the field presented on a silver platter;
- young specialists and prospective teachers, who want to make sure they see the forest for the trees, and haven't missed something significant.

As such, the target readership of this article is one that would most likely appreciate our simple take on cohomology and Hodge theory, via an approach that requires only linear algebra and graph theory.

In our approach, we have isolated the algebra from the topology to show that a large part of cohomology and Hodge theory is nothing more than the linear algebra of matrices satisfying $AB = 0$. For the remaining topological aspect, we cast our discussions entirely in terms of graphs as opposed to less-familiar topological objects like simplicial complexes. We believe that by putting these in a simple framework, we could facilitate the development of applications as well as communication with practitioners who may not otherwise see the utility of these notions.

Lastly, we write with a view towards readers whose main interests, like the author's, may lie in machine learning, matrix computations, numerical PDEs, optimization, statistics, or theory of computing, but have a casual interest in the topic and may perhaps want to explore potential applications in their respective fields.

The occasional whimsical section headings are inspired by [**11, 50, 51, 59, 60**].

2. Cohomology and Hodge theory for pedestrians

We will present in this section what we hope is the world's most elementary approach towards *cohomology* and *Hodge theory*, requiring only linear algebra.

2.1. Cohomology on a bumper sticker. Given two matrices $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ satisfying the property that

$$(2.1) \qquad\qquad AB = 0,$$

a property equivalent to

$$\mathrm{im}(B) \subseteq \ker(A),$$

the *cohomology group* with respect to A and B is the quotient vector space

$$\ker(A)/\mathrm{im}(B),$$

and its elements are called *cohomology classes*. The word 'group' here refers to the structure of $\ker(A)/\operatorname{im}(B)$ as an abelian group under addition.

We have fudged a bit because we haven't yet defined the matrices A and B. Cohomology really refers to a special case when A and B are certain matrices with topological meaning, as we will define in Section 3.

2.2. Harmonic representative. The definition in the previous section is plenty simple, provided the reader knows what a quotient vector space is, but can it be further simplified? For instance, can we do away with quotient spaces and equivalence classes[1] and define cohomology classes as actual vectors in $\mathbb{R}^n$?

Note that an element in $\ker(A)/\operatorname{im}(B)$ is a set of vectors

$$x + \operatorname{im}(B) := \{x + y \in \mathbb{R}^n : y \in \operatorname{im}(B)\}$$

for some $x \in \ker(A)$. We may avoid equivalence classes if we could choose an $x_H \in x + \operatorname{im}(B)$ in some unique way to represent the entire set. A standard way to do this is to pick x_H so that it is orthogonal to every other vector in $\operatorname{im}(B)$. Since $\operatorname{im}(B)^\perp = \ker(B^*)$, this is equivalent to requiring that $x_H \in \ker(B^*)$. Hence we should pick an $x_H \in \ker(A) \cap \ker(B^*)$. Such an x_H is called a *harmonic representative* of the cohomology class $x + \operatorname{im}(B)$.

The map that takes the cohomology class $x + \operatorname{im}(B)$ to its unique harmonic representative x_H gives a natural isomorphism of vector spaces

$$(2.2) \qquad \ker(A)/\operatorname{im}(B) \cong \ker(A) \cap \ker(B^*).$$

So we may redefine the cohomology group with respect to A and B to be the subspace $\ker(A) \cap \ker(B^*)$ of $\mathbb{R}^n$, and a cohomology class may now be regarded as an actual vector $x_H \in \ker(A) \cap \ker(B^*)$ (cf. Theorem A.3).

A word about our notation: B^* denotes the adjoint of the matrix B. Usually we will work over $\mathbb{R}$ with the standard l^2-inner product on our spaces and so $B^* = B^\mathsf{T}$. However we would like to allow for the possibility of working over $\mathbb{C}$ or with other inner products.

Numerical linear algebra interlude. For those familiar with numerical linear algebra, the way we choose a unique harmonic representative x_H to represent a cohomology class $x + \operatorname{im}(B)$ is similar to how we would impose uniqueness on a solution to a linear system of equations by requiring that it has minimum norm among all solutions [**34**, Section 5.5]. More specifically, the solutions to $Ax = b$ are given by $x_0 + \ker(A)$ where x_0 is any particular solution; we impose uniqueness by requiring that $x_0 \in \ker(A)^\perp = \operatorname{im}(A^*)$, which gives the minimum norm (or pseudoinverse) solution $x_0 = A^\dagger b$. The only difference above is that we deal with two matrices A and B instead of a single matrix A.

2.3. Hodge theory on one foot. We now explain why an element in $\ker(A) \cap \ker(B^*)$ is called 'harmonic'. Again assume that $AB = 0$, the *Hodge Laplacian* is the matrix

$$(2.3) \qquad A^*A + BB^* \in \mathbb{R}^{n \times n}.$$

It is an easy exercise (cf. Theorem A.2) to show that

$$(2.4) \qquad \ker(A^*A + BB^*) = \ker(A) \cap \ker(B^*).$$

---

[1]Practitioners tend to dislike working with equivalence classes of objects. One reason is that these are often tricky to implement in a computer program.

So the harmonic representative x_H that we constructed in Section 2.2 is a solution to the *Laplace equation*

$$(2.5) \qquad (A^*A + BB^*)x = 0.$$

Since solutions to the Laplace equation are called *harmonic* functions, this explains the name 'harmonic' representative.

With this observation, we see that we could also have defined the cohomology group (with respect to A and B) as the kernel of the Hodge Laplacian since

$$\ker(A)/\operatorname{im}(B) \cong \ker(A^*A + BB^*).$$

It is also an easy exercise (cf. Theorem A.2) to show that one has a *Hodge decomposition*, an orthogonal direct sum decomposition

$$(2.6) \qquad \mathbb{R}^n = \operatorname{im}(A^*) \oplus \ker(A^*A + BB^*) \oplus \operatorname{im}(B).$$

In other words, whenever $AB = 0$, every $x \in \mathbb{R}^n$ can be decomposed uniquely as

$$x = A^*w + x_H + Bv, \qquad \langle A^*w, x_H \rangle = \langle x_H, Bv \rangle = \langle A^*w, Bv \rangle = 0,$$

for some $v \in \mathbb{R}^p$ and $w \in \mathbb{R}^m$.

Recall the well-known decompositions (sometimes called the Fredholm alternative) associated with the four fundamental subspaces [64] of a matrix $A \in \mathbb{R}^{m \times n}$,

$$(2.7) \qquad \mathbb{R}^n = \ker(A) \oplus \operatorname{im}(A^*), \quad \mathbb{R}^m = \ker(A^*) \oplus \operatorname{im}(A).$$

The Hodge decomposition (2.6) may be viewed as an analogue of (2.7) for a pair of matrices satisfying $AB = 0$. In fact, combining (2.6) and (2.7), we obtain

$$(2.8) \qquad \mathbb{R}^n = \overbrace{\operatorname{im}(A^*) \oplus \underbrace{\ker(A^*A + BB^*) \oplus \operatorname{im}(B)}_{\ker(A)}}^{\ker(B^*)}.$$

The intersection of $\ker(A)$ and $\ker(B^*)$ gives $\ker(A^*A + BB^*)$, confirming (2.4). Since $A^*A + BB^*$ is Hermitian, it also follows that

$$(2.9) \qquad \operatorname{im}(A^*A + BB^*) = \operatorname{im}(A^*) \oplus \operatorname{im}(B).$$

For the special case when A is an arbitrary matrix and $B = 0$, the Hodge decomposition (2.6) becomes

$$(2.10) \qquad \mathbb{R}^n = \operatorname{im}(A) \oplus \ker(A^*A),$$

which may also be deduced directly from the Fredholm alternative (2.7) since

$$(2.11) \qquad \ker(A^*A) = \ker(A).$$

Numerical linear algebra interlude. To paint an analogy like that in the last paragraph of Section 2.2, our characterization of cohomology classes as solutions to the Laplace equation (2.5) is similar to the characterization of solutions to a least squares problem $\min_{x \in \mathbb{R}^n} \|Ax - b\|$ as solutions to its normal equation $A^*Ax = A^*b$ [34, Section 6.3]. Again the only difference is that here we deal with two matrices instead of just one.

2.4. Terminologies. One obstacle that the (impatient) beginner often faced when learning cohomology is the considerable number of scary-sounding terminologies that we have by-and-large avoided in the treatment above.

In Table 1, we summarize some commonly used terminologies for objects in Sections 2.1, 2.2, and 2.3. Their precise meanings will be given in Sections 3 and 4, with an updated version of this table appearing as Table 3. As the reader can

TABLE 1. Topological jargons (first pass)

NAME	MEANING
coboundary maps	$A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$
cochains	elements of $\mathbb{R}^n$
cochain complex	$\mathbb{R}^p \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$
cocycles	elements of $\ker(A)$
coboundaries	elements of $\operatorname{im}(B)$
cohomology classes	elements of $\ker(A)/\operatorname{im}(B)$
harmonic cochains	elements of $\ker(A^*A + BB^*)$
Betti numbers	$\dim \ker(A^*A + BB^*)$
Hodge Laplacians	$A^*A + BB^* \in \mathbb{R}^{n \times n}$
x is closed	$Ax = 0$
x is exact	$x = Bv$ for some $v \in \mathbb{R}^p$
x is coclosed	$B^*x = 0$
x is coexact	$x = A^*w$ for some $w \in \mathbb{R}^m$
x is harmonic	$(A^*A + BB^*)x = 0$

see, there is some amount of redundancy in these terminologies; e.g., saying that a cochain is exact is the same as saying that it is a coboundary. This can sometimes add to the confusion for a beginner. It is easiest to just remember equations and disregard jargons. When people say things like 'a cochain is harmonic if and only if it is closed and coclosed,' they are just verbalizing (2.4).

In summary, we saw *three different ways of defining cohomology*: If A and B are matrices satisfying $AB = 0$, then the cohomology group with respect to A and B may be taken to be any one of the following,

$$(2.12) \qquad \ker(A)/\operatorname{im}(B), \qquad \ker(A) \cap \ker(B^*), \qquad \ker(A^*A + BB^*).$$

For readers who have heard of the term *homology*, that can be defined just by taking adjoints. Note that if $AB = 0$, then $B^*A^* = 0$ and we may let B^* and A^* play the role of A and B respectively. The homology group with respect to A and B may then be taken to be any one of the following,

$$(2.13) \qquad \ker(B^*)/\operatorname{im}(A^*), \qquad \ker(B^*) \cap \ker(A), \qquad \ker(BB^* + A^*A).$$

As we can see, the last two spaces in (2.12) and (2.13) are identical, i.e., there is no difference between cohomology and homology in our context (but see Section 5.1 for caveats).

3. Coboundary operators and Hodge Laplacians on graphs

Observe that everything in Section 2 relies only on the assumption that $AB = 0$. We haven't been entirely forthright: What we have discussed up to this point is really the linear algebra of operators satisfying $AB = 0$ — the 'algebraic half' of of cohomology and Hodge theory. There is also a 'topological half' that is just one step away — they refer to the special case when A and B are the so-called coboundary operators. This special case allows us to attach topological meanings to the objects in Section 2.

Just like the last section requires nothing more than elementary linear algebra, this section requires nothing more than elementary graph theory. We will discuss simplicial complexes (family of subsets of vertices), cochains (functions on a graph), and coboundary operators (operators on functions on a graph) — all in the context of the simplest type of graphs: undirected, unweighted, no loops, no multiple edges.

3.1. Graphs. Let $G = (V, E)$ be an undirected graph where $V := \{1, \ldots, n\}$ is a finite set of vertices and $E \subseteq \binom{V}{2}$ is the set[2] of edges. Note that once we have specified G, we automatically get *cliques* of higher order — for example, the set of triangles or 3-*cliques* $T \subseteq \binom{V}{3}$ is defined by

$$\{i, j, k\} \in T \quad \text{iff} \quad \{i, j\}, \{i, k\}, \{j, k\} \in E.$$

More generally the set of k-*cliques* $K_k(G) \subseteq \binom{V}{k}$ is defined by

$$\{i_1, \ldots, i_k\} \in K_k(G) \quad \text{iff} \quad \{i_p, i_q\} \in E \text{ for all } 1 \le p < q \le k,$$

i.e., all pairs of vertices in $\{i_1, \ldots, i_k\}$ are in E.

In particular we have

$$K_1(G) = V, \quad K_2(G) = E, \quad K_3(G) = T.$$

Note that once we have specified V and E, all $K_k(G)$ for $k \ge 3$ would have been uniquely determined.

In topological parlance, a nonempty family K of finite subsets of a set V is called a *simplicial complex* (more accurately, an abstract simplicial complex) if for any set S in K, every $S' \subseteq S$ also belongs to K. Evidently the set comprising all cliques of a graph G,

$$K(G) := \bigcup_{k=1}^{\omega(G)} K_k(G),$$

is a simplicial complex and is called the *clique complex* of the graph G. The *clique number* $\omega(G)$ is the number of vertices in a largest clique of G.

There are abstract simplicial complexes that are not clique complexes of graphs. For example, we may just exclude cliques of larger sizes — $\bigcup_{k=1}^{m} K_k(G)$ is still an abstract simplicial complex for any $m = 3, \ldots, \omega(G) - 1$, but it would not in general be a clique complex of a graph.

3.2. Functions on a graph. Given a graph $G = (V, E)$, we may define real-valued functions on its vertices $f : V \to \mathbb{R}$. We may also define real-valued functions on E and T and $K_k(G)$ in general but we shall require them to be *alternating*. By an alternating function on E, we mean one of the form $X : V \times V \to \mathbb{R}$ where

$$X(i, j) = -X(j, i)$$

---

[2]Henceforth $\binom{V}{k}$ denotes the set of all k-element subsets of V. In particular E is not a multiset since our graphs have no loops nor multiple edges.

for all $\{i, j\} \in E$, and

$$X(i, j) = 0$$

for all $\{i, j\} \notin E$. An alternating function on T is one of the form $\Phi : V \times V \times V \to \mathbb{R}$ where

$$\Phi(i, j, k) = \Phi(j, k, i) = \Phi(k, i, j) = -\Phi(j, i, k) = -\Phi(i, k, j) = -\Phi(k, j, i)$$

for all $\{i, j, k\} \in T$, and

$$\Phi(i, j, k) = 0$$

for all $\{i, j, k\} \notin T$. More generally, an alternating function is one where permutation of its arguments has the effect of multiplying by the sign of the permutation, as we will see in (4.1).

In topological parlance, the functions f, X, Φ are called 0-, 1-, 2-*cochains*. These are discrete analogues of *differential forms* on manifolds [**67**]. Those who prefer to view them as such often refer to cochains as *discrete differential forms* [**27, 39**] or *Whitney forms* [**12, 68**] and in which case, f, X, Φ are 0-, 1-, 2-*forms* on G.

Observe that a 1-cochain X is completely specified by the values it takes on the set $\{(i, j) : i < j\}$ and a 2-cochain Φ is completely specified by the values it takes on the set $\{(i, j, k) : i < j < k\}$. We may equip the spaces of cochains with inner products, for example, as weighted sums

$$(3.1) \qquad \langle f, g \rangle_V = \sum_{i=1}^{n} w_i f(i) g(i), \qquad \langle X, Y \rangle_E = \sum_{i<j} w_{ij} X(i, j) Y(i, j),$$

$$\langle \Phi, \Psi \rangle_T = \sum_{i<j<k} w_{ijk} \Phi(i, j, k) \Psi(i, j, k),$$

where the weights w_i, w_{ij}, w_{ijk} are given by any positive values invariant under arbitrary permutation of indices. When they take the constant value 1, we call it the *standard l^2-inner product*. By summing only over the sets[3] $\{(i, j) : i < j\}$ and $\{(i, j, k) : i < j < k\}$, we count each edge or triangle exactly once in the inner products.

We will denote the Hilbert spaces of 0-, 1-, and 2-cochains as $L^2(V)$, $L^2_\wedge(E)$, $L^2_\wedge(T)$ respectively. The subscript $\wedge$ is intended to indicate 'alternating'. Note that $L^2_\wedge(V) = L^2(V)$ since for a function of one argument, being alternating is a vacuous property. We set $L^2_\wedge(\varnothing) := \{0\}$ by convention.

The elements of $L^2_\wedge(E)$ (i.e., 1-cochains) are well-known in graph theory, often called *edge flows*. While the graphs in this article are always undirected and unweighted, a directed graph is simply one equipped with a choice of edge flow $X \in L^2_\wedge(E)$ — an undirected edge $\{i, j\} \in E$ becomes a directed edge (i, j) if $X(i, j) > 0$ or (j, i) if $X(i, j) < 0$; and the magnitude of $X(i, j)$ may be taken as the weight of that directed edge. So $L^2_\wedge(E)$ encodes all weighted directed graphs that have the same underlying undirected graph structure.

3.3. Operators on functions on a graph. We will consider the graph-theoretic analogues of grad, curl, div in multivariate calculus. The *gradient* is the linear operator $\mathrm{grad} : L^2(V) \to L^2_\wedge(E)$ defined by

$$(\mathrm{grad}\, f)(i, j) = f(j) - f(i)$$

---

[3] Our choice is arbitrary; any set that includes each edge or triangle exactly once would also serve the purpose. Each such choice corresponds to a choice of orientation on elements of E or T.

for all $\{i, j\} \in E$ and zero otherwise. The *curl* is the linear operator curl : $L_\wedge^2(E) \to$ $L_\wedge^2(T)$ defined by

$$(\operatorname{curl} X)(i, j, k) = X(i, j) + X(j, k) + X(k, i)$$

for all $\{i, j, k\} \in T$ and zero otherwise. The *divergence* is the linear operator div : $L_\wedge^2(E) \to L^2(V)$ defined by

$$(\operatorname{div} X)(i) = \sum_{j=1}^{n} \frac{w_{ij}}{w_i} X(i, j)$$

for all $i \in V$.

Using these, we may construct other linear operators, notably the well-known *graph Laplacian*, the operator $\Delta_0 : L^2(V) \to L^2(V)$ defined by

$$\Delta_0 = - \operatorname{div} \operatorname{grad},$$

which is a graph-theoretic analogue of the Laplace operator (cf. Lemma B.3). Less well-known is the *graph Helmholtzian*, the operator $\Delta_1 : L_\wedge^2(E) \to L_\wedge^2(E)$ defined by

$$\Delta_1 = - \operatorname{grad} \operatorname{div} + \operatorname{curl}^* \operatorname{curl},$$

which is a graph-theoretic analogue of the Helmholtz or vector Laplace operator. It is straightforward to derive an expression for the adjoint of the curl operator, $\operatorname{curl}^* : L_\wedge^2(T) \to L_\wedge^2(E)$ is given by

$$(\operatorname{curl}^* \Phi)(i, j) = \sum_{k=1}^{n} \frac{w_{ijk}}{w_{ij}} \Phi(i, j, k)$$

for all $\{i, j\} \in E$ and zero otherwise (cf. Lemma B.2).

The gradient and curl are special cases of *coboundary operators*, discrete analogues of *exterior derivatives*, while the graph Laplacian and Helmholtzian are special cases of *Hodge Laplacians*.

The matrices A and B that we left unspecified in Section 2 are coboundary operators. It is easy to see that the composition

$$(3.2) \qquad\qquad\qquad \operatorname{curl} \operatorname{grad} = 0$$

and so setting $A = \operatorname{curl}$ and $B = \operatorname{grad}$ gives us (2.1) (cf. Theorem B.4).

Note that divergence and gradient are negative adjoints of each other:

$$(3.3) \qquad\qquad\qquad \operatorname{div} = - \operatorname{grad}^*,$$

(cf. Lemma B.1). With this we get $\Delta_1 = A^* A + B B^*$ as in (2.3).

If the inner products on $L^2(V)$ and $L_\wedge^2(E)$ are taken to be the standard l^2-inner products, then (3.3) gives $\Delta_0 = B^* B = B^\mathsf{T} B$, a well-known expression of the graph Laplacian in terms of vertex-edge incidence matrix B. The operators

$$\operatorname{grad}^* \operatorname{grad} : L^2(V) \to L^2(V) \quad \text{and} \quad \operatorname{curl}^* \operatorname{curl} : L_\wedge^2(E) \to L_\wedge^2(E)$$

are sometimes called the *vertex Laplacian* and *edge Laplacian* respectively. The vertex Laplacian is of course just the usual graph Laplacian but note that the edge Laplacian is not the same as the graph Helmholtzian.

Physics interlude. Take the standard l^2-inner products on $L^2(V)$ and $L^2_\wedge(E)$, the divergence of an edge flow at a vertex $i \in V$ may be interpreted as the *netflow*,

$$(3.4) \qquad (\operatorname{div} X)(i) = (\operatorname{inflow} X)(i) - (\operatorname{outflow} X)(i),$$

where *inflow* and *outflow* are defined respectively for any $X \in L^2_\wedge(E)$ and any $i \in V$ as

$$(\operatorname{inflow} X)(i) = \sum\nolimits_{j:X(i,j)>0} X(i,j), \quad (\operatorname{outflow} X)(i) = \sum\nolimits_{j:X(i,j)<0} X(i,j).$$

Sometimes the terms *incoming flux*, *outgoing flux*, *total flux* are used instead of inflow, outflow, net flow. Figure 2 shows two *divergence-free* edge flows, i.e., inflow equals outflow at every vertex.

Let $X \in L^2_\wedge(E)$. A vertex $i \in V$ is called a *sink* of X if $X(i,j) > 0$ for every neighbor $\{i,j\} \in E$ of i. Likewise a vertex $i \in V$ is called a *source* of X if $X(i,j) < 0$ for every neighbor $\{i,j\} \in E$ of i. In general, an edge flow may not have any source or sink[4] but if it can be written as

$$(3.5) \qquad\qquad\qquad\qquad X = -\operatorname{grad} f$$

for some $f \in L^2(V)$, often called a *potential* function, then X will have the property of flowing from sources (local maxima of f) to sinks (local minima of f). See Figure 8B for an example of an edge flow given by the gradient of a potential function, flowing from two sources into two sinks.

EXAMPLE 3.1. We highlight a common pitfall regarding curl on a graph. Consider C_3 and C_4, the *cycle graphs* on three and four vertices in Figure 1.

FIGURE 1. Cycle graphs C_3 (*left*) and C_4 (*right*).

Number the vertices and consider the edge flows in Figure 2. What are the values of their curl? For the one on C_3, the answer is $2 + 2 + 2 = 6$ as expected. But the answer for the one on C_4 is not $2 + 2 + 2 + 2 = 8$, it is in fact 0.
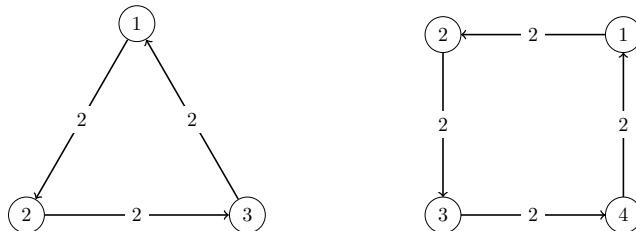

FIGURE 2. Edge flows on C_3 (*left*) and C_4 (*right*).

---

[4]There is an alternative convention that defines $i \in V$ to be a source (resp. sink) of X as long as $\operatorname{div} X(i) > 0$ (resp. $\operatorname{div} X(i) < 0$) but our definition is much more restrictive.

The second answer may not agree with a physicist's intuitive idea of curl and is a departure from what one would expect in the continuous case. However it is what follows from the definition. Let $X \in L^2_\wedge(E(C_3))$ denote the edge flow on C_3 in Figure 2. It is given by

$$X(1,2) = X(2,3) = X(3,1) = 2 = -X(2,1) = -X(3,2) = -X(1,3),$$

and the curl evaluated at $\{1,2,3\} \in T(C_3)$ is by definition indeed

$$(\operatorname{curl} X)(1,2,3) = X(1,2) + X(2,3) + X(3,1) = 6.$$

On the other hand C_4 has no 3-cliques and so $T(C_4) = \varnothing$. By convention $L^2_\wedge(\varnothing) = \{0\}$. Hence curl : $L^2_\wedge(E(C_4)) \to L^2_\wedge(T(C_4))$ must have curl $X = 0$ for all $X \in L^2_\wedge(E(C_4))$ and in particular for the edge flow on the right of Figure 2.

TABLE 2. Electrodynamics/fluid dynamics jargons

NAME	MEANING	ALTERNATE NAME(S)
divergence-free	element of ker(div)	solenoidal
curl-free	element of ker(curl)	irrotational
vorticity	element of im(curl*)	vector potential
conservative	element of im(grad)	potential flow
harmonic	element of ker(Δ_1)	
anharmonic	element of im(Δ_1)	
scalar field	element of $L^2(V)$	scalar potential
vector field	element of $L^2_\wedge(E)$	

3.4. Helmholtz decomposition for graphs. The graph Laplacian $\Delta_0 : L^2(V) \to L^2(V)$,

$$\Delta_0 = -\operatorname{div}\operatorname{grad} = \operatorname{grad}^* \operatorname{grad},$$

particularly under the context of spectral graph theory [21, 63], has been an enormously useful construct, with great impact on many areas. We have nothing more to add except to remark that the Hodge decomposition of Δ_0 is given by (2.10),

$$L^2(V) = \ker(\Delta_0) \oplus \operatorname{im}(\operatorname{div}).$$

Recall from (2.11) that $\ker(\Delta_0) = \ker(\operatorname{grad})$. Since $\operatorname{grad} f = 0$ iff f is piecewise constant, i.e., constant on each connected component of G, the number $\beta_0(G) = \dim \ker(\Delta_0)$ counts the number of connected component of G — a well known fact in graph theory.

The Hodge decomposition for the graph Helmholtzian $\Delta_1 : L^2_\wedge(E) \to L^2_\wedge(E)$,

$$\Delta_1 = -\operatorname{grad}\operatorname{div} + \operatorname{curl}^*\operatorname{curl} = \operatorname{grad}\operatorname{grad}^* + \operatorname{curl}^*\operatorname{curl}.$$

is called the *Helmholtz decomposition*. It says that the space of edge flows admits an orthogonal decomposition into subspaces

$$(3.6) \qquad L^2_\wedge(E) = \overbrace{\operatorname{im}(\operatorname{curl}^*) \oplus \underbrace{\ker(\Delta_1)}_{} }^{\ker(\operatorname{div})} \underbrace{\oplus \operatorname{im}(\operatorname{grad})}_{\ker(\operatorname{curl})},$$

and moreover the three subspaces are related via

$$(3.7) \qquad \ker(\Delta_1) = \ker(\mathrm{curl}) \cap \ker(\mathrm{div}), \qquad \mathrm{im}(\Delta_1) = \mathrm{im}(\mathrm{curl}^*) \oplus \mathrm{im}(\mathrm{grad}).$$

In particular, the first equation is a discrete analogue of the statement "a vector field is curl-free and divergence-free if and only if it is a harmonic vector field."

There is nothing really special here — as we saw in Section 2.3, any matrices A and B satisfying $AB = 0$ would give such a decomposition: (3.6) and (3.7) are indeed just (2.6), (2.4), and (2.9) where $A = \mathrm{curl}$ and $B = \mathrm{grad}$. This is however a case that yields the most interesting applications (see Section 6 and [**15, 43**]).

EXAMPLE 3.2 (Beautiful mind problem on graphs). This is a discrete analogue of a problem[5] that appeared in a blockbuster movie: Let $G = (V, E)$ be a graph. If $X \in L^2_\wedge(E)$ is curl-free, then is it true that X is a gradient? In other words, if $X \in \ker(\mathrm{curl})$, must it also be in $\mathrm{im}(\mathrm{grad})$? Clearly the converse always holds by (3.2) but from (3.6), we know that

$$(3.8) \qquad \qquad \ker(\mathrm{curl}) = \ker(\Delta_1) \oplus \mathrm{im}(\mathrm{grad})$$

and so it is not surprising that the answer is generally no. We would like to describe a family of graphs for which the answer is yes.
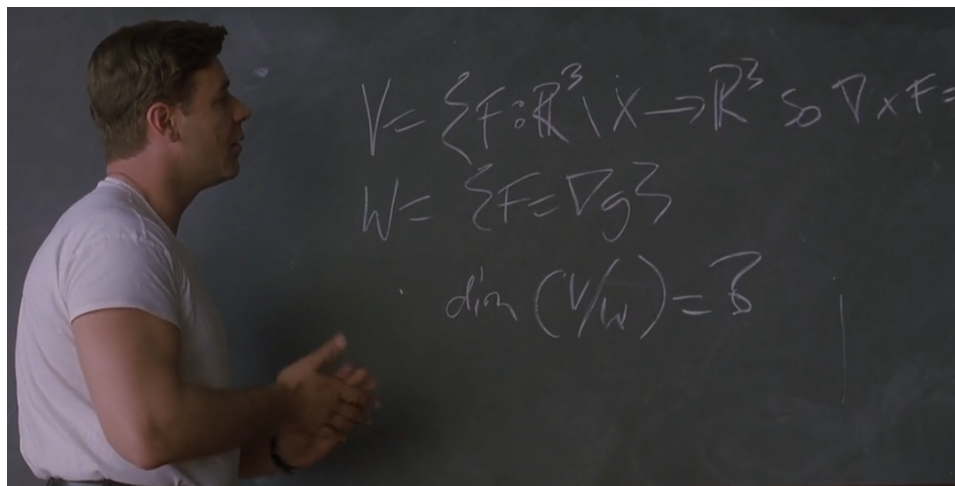

FIGURE 3. Problem from *A Beautiful Mind*: $V = \{F : \mathbb{R}^3 \setminus X \to \mathbb{R}^3 \text{ so } \nabla \times F = 0\}$, $W = \{F = \nabla g\}$, $\dim(V/W) = ?$

The edge flow $X \in L^2_\wedge(E(C_4))$ on the right of Figure 2 is an example of one that is curl-free but not a gradient. It is trivially curl-free since $T(C_4) = \varnothing$. It is not a gradient since if $X = \mathrm{grad}\, f$, then

$$f(2) - f(1) = 2, \quad f(3) - f(2) = 2, \quad f(4) - f(3) = 2, \quad f(1) - f(4) = 2,$$

and summing them gives '0 = 8' — a contradiction. Note that X is also divergence-free by (3.4) since inflow X = outflow X. It is therefore harmonic by (3.7), i.e., $X \in \ker(\Delta_1)$ as expected.

---

[5]Due to Dave Bayer [**13**], now a ubiquitous homework problem in many multivariate calculus courses thanks to him and the movie. See Figure 3.

Every divergence-free edge flow on C_4 must be of the same form as X, taking constant value on all edges or otherwise we would not have inflow X = outflow X. Since all edge flows on C_4 are automatically curl-free, $\ker(\Delta_1) = \ker(\mathrm{div})$ and is given by the set of all constant multiples of X. The number

$$\beta_1(G) = \dim \ker(\Delta_1)$$

counts the number of '1-dimensional holes' of G and in this case we see that indeed $\beta_1(C_4) = 1$. To be a bit more precise, the '1-dimensional holes' are the regions that remain uncovered after the cliques are filled in.

We now turn our attention to the contrasting case of C_3. Looking at Figure 1, it may seem that C_3 also has a '1-dimensional hole' as in C_4 but this is a fallacy — holes bounded by triangles are not regarded as holes in our framework.

For C_3 it is in fact true that every curl-free edge flow is a gradient. To see this, note that as in the case of C_4, any divergence-free $X \in L^2_\wedge(E(C_3))$ must be constant on all edges and so

$$(\mathrm{curl}\, X)(1,2,3) = X(1,2) + X(2,3) + X(3,1) = c + c + c = 3c,$$

for some $c \in \mathbb{R}$. If a divergence-free X is also curl-free, then $c = 0$ and so $X = 0$. Hence for C_3, $\ker(\Delta_1) = \{0\}$ by (3.8) and $\ker(\mathrm{curl}) = \mathrm{im}(\mathrm{grad})$ by (3.7). It also follows that $\beta_1(C_3) = 0$ and so C_3 has no '1-dimensional hole'.

What we have illustrated with C_3 and C_4 extends to any arbitrary graph. A moment's thought would reveal that the property $\beta_1(G) = 0$ is satisfied by any *chordal graph*, i.e., one for which every cycle subgraph of four or more vertices has a *chord*, an edge that connects two vertices of the cycle subgraph but that is not part of the cycle subgraph. Equivalently, a chordal graph is one where every chordless cycle subgraph is C_3.

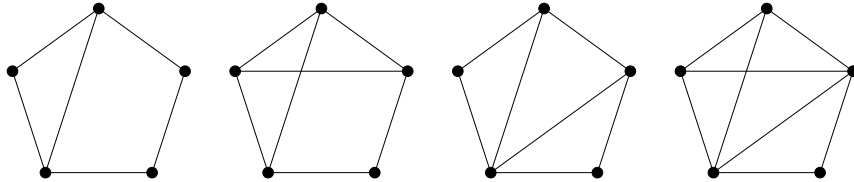
FIGURE 4. Left two graphs: not chordal. Right two graphs: chordal.

There are however non-chordal graphs with $\beta_1(G) = 0$. Take for example the *wheel graph* W_n, formed by joining all $n - 1$ vertices of the cycle graph C_{n-1} to an extraneous nth vertex, $n \geq 4$. It is clear that $\beta_1(W_n) = 0$ for all $n \geq 4$ but that only W_4 is a chordal graph. For the wheel graphs W_5, W_6, W_7 in Figure 5, their cycle subgraphs C_4, C_5, C_6 are evidently all chordless.
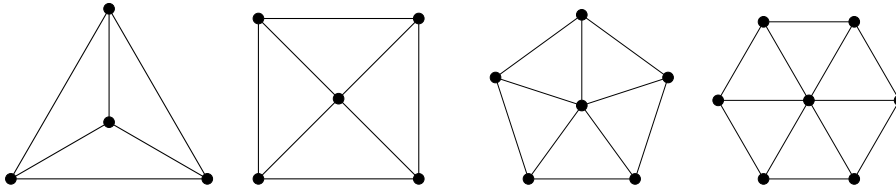

FIGURE 5. Wheel graphs W_4, W_5, W_6, W_7 (left to right).

4. Higher order

We expect the case of alternating functions on edges, i.e., $k = 1$, discussed in Section 3 to be the most useful in applications. However for completeness and since it is no more difficult to generalize to $k > 1$, we provide the analogue of Section 3 for arbitrary k here.

4.1. Higher-order cochains. Let $K(G)$ be the clique complex of a graph $G = (V, E)$ as defined in Section 3.1. We will write $K_k = K_k(G)$ for simplicity.

A k-*cochain* (or k-form) is an alternating function on K_{k+1}, or more specifically, $f : V \times \cdots \times V \to \mathbb{R}$ where

$$(4.1) \qquad f(i_{\sigma(0)}, \ldots, i_{\sigma(k)}) = \operatorname{sgn}(\sigma) f(i_0, \ldots, i_k)$$

for all $\{i_0, \ldots, i_k\} \in K_{k+1}$ and all $\sigma \in \mathfrak{S}_{k+1}$, the symmetric group of permutations on $\{0, \ldots, k\}$. We set $f(i_0, \ldots, i_k) = 0$ if $\{i_0, \ldots, i_k\} \notin K_{k+1}$.

Again, we may put an inner product on k-cochains,

$$\langle f, g \rangle = \sum_{i_0 < \cdots < i_k} w_{i_0 \cdots i_k} f(i_0, \ldots, i_k) g(i_0, \ldots, i_k),$$

with any positive weights satisfying $w_{i_{\sigma(0)} \cdots i_{\sigma(k)}} = w_{i_0 \cdots i_k}$ for all $\sigma \in \mathfrak{S}_{k+1}$.

We denote the resulting Hilbert space by $L^2_\wedge(K_{k+1})$. This is a subspace of $L^2(\bigwedge^{k+1} V)$, the space of alternating functions with $k + 1$ arguments in V. Clearly,

$$\dim L^2_\wedge(K_{k+1}) = \#K_{k+1}.$$

A word of caution regarding the terminology: a k-cochain is a function on a $(k + 1)$-clique and has $k + 1$ arguments. The reason is due to the different naming conventions — a $(k+1)$-clique in graph theory is called a k-simplex in topology. In topological lingo, a vertex is a 0-simplex, an edge a 1-simplex, a triangle a 2-simplex, a tetrahedron a 3-simplex.

4.2. Higher-order coboundary operators. The k-*coboundary operators* $\delta_k : L^2_\wedge(K_k) \to L^2_\wedge(K_{k+1})$ are defined by

$$(4.2) \qquad (\delta_k f)(i_0, \ldots, i_{k+1}) = \sum_{j=0}^{k+1} (-1)^j f(i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_{k+1}),$$

for $k = 0, 1, 2, \ldots$. Readers familiar with differential forms may find it illuminating to think of coboundary operators as discrete analogues of exterior derivatives. Note that f is a function with $k+1$ arguments but $\delta_k f$ is a function with $k+2$ arguments. A convenient oft-used notation is to put a carat over the omitted argument

$$(4.3) \qquad f(i_0, \ldots, \widehat{i}_j, \ldots, i_{k+1}) := f(i_0, \ldots, i_{j-1}, i_{j+1}, \ldots, i_{k+1}).$$

The crucial relation $AB = 0$ in Section 2 is in fact

$$(4.4) \qquad \delta_k \delta_{k-1} = 0,$$

which is straightforward to verify using (4.2) (cf. Theorem B.4). This generalizes (3.2) and is sometimes called the *fundamental theorem of topology*:

> The boundary of a boundary is empty.

The reader is encouraged to envision any common geometric objects and see this for himself. The message expressed by (4.4) is really the dual statement:

> The coboundary of a coboundary is zero.

As in Section 2.1, (4.4) is equivalent to saying $\operatorname{im}(\delta_{k-1})$ is a subspace of $\ker(\delta_k)$. We define the kth *cohomology group* of G to be the quotient vector space

$$H^k(G) = \ker(\delta_k)/\operatorname{im}(\delta_{k-1}),$$

for $k = 1, 2, \ldots, \omega(G) - 1$.

To keep track of the coboundary operators, it is customary to assemble them into a sequence of maps written in the form

$$L_\wedge^2(K_0) \xrightarrow{\delta_0} L_\wedge^2(K_1) \xrightarrow{\delta_1} \cdots \xrightarrow{\delta_{k-1}} L_\wedge^2(K_k) \xrightarrow{\delta_k} L_\wedge^2(K_{k+1}) \xrightarrow{\delta_{k+1}} \cdots \xrightarrow{\delta_\omega} L_\wedge^2(K_\omega).$$

This sequence is called a *cochain complex*.

For $k = 1$, we get $\delta_0 = \mathrm{grad}$, $\delta_1 = \mathrm{curl}$, and the first two terms of the cochain complex are

$$L^2(V) \xrightarrow{\mathrm{grad}} L_\wedge^2(E) \xrightarrow{\mathrm{curl}} L_\wedge^2(T).$$

4.3. Hodge theory. The *Hodge k-Laplacian* $\Delta_k : L_\wedge^2(K_k) \to L_\wedge^2(K_k)$ is defined as

$$\Delta_k = \delta_{k-1}\delta_{k-1}^* + \delta_k^*\delta_k.$$

We call $f \in L_\wedge^2(K_k)$ a *harmonic k-cochain* if it satisfies the Laplace equation

$$\Delta_k f = 0.$$

Applying the results in Section 2.3 with $A = \delta_k$ and $B = \delta_{k-1}$, we obtain the unique representation of cohomology classes as harmonic cochains

$$H^k(G) = \ker(\delta_k)/\operatorname{im}(\delta_{k-1}) \cong \ker(\delta_k) \cap \ker(\delta_{k-1}^*) = \ker(\Delta_k),$$

as well as the Hodge decomposition

$$L_\wedge^2(K_k) = \overbrace{\operatorname{im}(\delta_k^*) \oplus \underbrace{\ker(\Delta_k)}_{} \oplus \operatorname{im}(\delta_{k-1})}^{\ker(\delta_{k-1}^*)},$$

and the relation

$$\operatorname{im}(\Delta_1) = \operatorname{im}(\delta_k^*) \oplus \operatorname{im}(\delta_{k-1}).$$

EXAMPLE 4.1 (Graph isomorphism). Two undirected graphs G and H on n vertices are said to be *isomorphic* if they are essentially the same graph up to relabeling of vertices. The *graph isomorphism problem*, an open[6] problem in computer science, asks whether there is a polynomial-time algorithm for deciding if two given graphs are isomorphic [5]. Clearly two isomorphic graphs must be *isospectral* in the sense that the eigenvalues (ordered and counted with multiplicities) of their graph Laplacians are equal,

$$\lambda_i(\Delta_0(G)) = \lambda_i(\Delta_0(H)), \quad i = 1, \ldots, n,$$

a condition that can be checked in polynomial time. Not surprisingly, the converse is not true, or we would have been able to determine graph isomorphism in polynomial time. We should mention that there are several definitions of isospectral graphs, in terms of the adjacency matrix, graph Laplacian, normalized Laplacian, signless Laplacian, etc. See [**14, 36**] for many interesting examples.

---

[6]Perhaps not for much longer — a significant breakthrough has been announced [**6**] as this article is going to press.

The reader may perhaps wonder what happens if we impose the stronger requirement that the eigenvalues of all their higher-order Hodge k-Laplacians be equal as well?

$$\lambda_i(\Delta_k(G)) = \lambda_i(\Delta_k(H)), \quad i = 1, \ldots, n, \ k = 0, \ldots, m.$$

For any $m \geq 1$, these indeed give a stronger set of sufficient conditions that can be checked in polynomial time. For example, the eigenvalues of Δ_0 for the two graphs in Figure 6 are $0, 0.76, 2, 3, 3, 5.24$ (all numbers rounded to two decimal figures). On the other hand, the eigenvalues of Δ_1 are $0, 0.76, 2, 3, 3, 3, 5.24$ for the graph on the left and $0, 0, 0.76, 2, 3, 3, 5.24$ for the graph on the right, allowing us to conclude that they are not isomorphic. These calculations are included in Section C.

FIGURE 6. These graphs have isospectral Laplacians (Hodge 0-Laplacians) but not Helmholtzians (Hodge 1-Laplacians).

Non-isomorphic graphs can nevertheless have isospectral Hodge Laplacians of all order. The two graphs in Figure 7 are clearly non-isomorphic. Neither contains cliques of order higher than two, so their Hodge k-Laplacians are zero for all $k > 2$. It is straightforward to check that the first three Hodge Laplacians Δ_0, Δ_1, Δ_2, of both graphs are isospectral (cf. Section C).
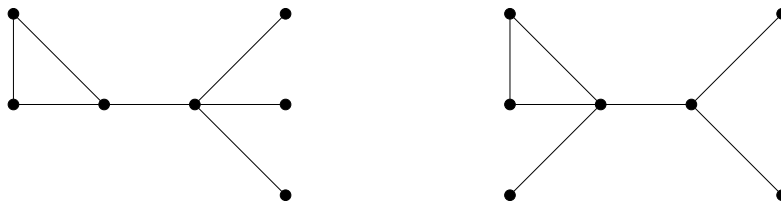

FIGURE 7. Non-isormorphic graphs with isospectral Hodge k-Laplacians for all $k = 0, 1, 2, \ldots$.

5. Further information

5.1. Topological caveats. The way we defined cohomology in Section 2.1 is more or less standard. The only simplification is that we had worked over a field. The notion of cohomology in topology works more generally over arbitrary rings where this simple linear algebraic approach falls short, but not by much — all we need is to be willing to work with modules over rings instead of modules over fields, i.e., vector spaces. Unlike a vector space, a module may not have a basis and we may not necessarily be able to represent linear maps by matrices, a relatively small price to pay.

TABLE 3. Topological jargons (second pass)

NAME	MEANING
coboundary maps	$\delta_k : L^2_\wedge(K_k) \to L^2_\wedge(K_{k+1})$
cochains	elements of $L^2_\wedge(K_k)$
cochain complex	$\cdots \longrightarrow L^2_\wedge(K_{k-1}) \xrightarrow{\delta_{k-1}} L^2_\wedge(K_k) \xrightarrow{\delta_k} L^2_\wedge(K_{k+1}) \longrightarrow \cdots$
cocycles	elements of $\ker(\delta_k)$
coboundaries	elements of $\operatorname{im}(\delta_{k-1})$
cohomology classes	elements of $\ker(\delta_k)/\operatorname{im}(\delta_{k-1})$
harmonic cochains	elements of $\ker(\Delta_k)$
Betti numbers	$\dim \ker(\Delta_k)$
Hodge Laplacians	$\Delta_k = \delta_{k-1}\delta^*_{k-1} + \delta^*_k \delta_k$
f is closed	$\delta_k f = 0$
f is exact	$f = \delta_{k-1}g$ for some $g \in L^2_\wedge(K_{k-1})$
f is coclosed	$\delta^*_{k-1}f = 0$
f is coexact	$f = \delta^*_k h$ for some $h \in L^2_\wedge(K_{k+1})$
f is harmonic	$\Delta_k f = 0$

However the further simplifications in Sections 2.2 and 2.3 to avoid quotient spaces only hold when we have a field of characteristic zero (we chose $\mathbb{R}$). For example, if instead of $\mathbb{R}$, we had the field $\mathbb{F}_2$ of two elements with binary arithmetic (or indeed any field of positive characteristic), then we can no longer define inner products and statements like $\ker(B)^\perp = \operatorname{im}(B^*)$ make no sense. While the adjoint of a matrix may still be defined without reference to an inner product, statements like $\ker(A^*A) = \ker(A)$ are manifestly false in positive characteristic.

We saw in Section 2.4 that in the way we presented things, there is no difference between cohomology and homology. This is an artifact of working over a field. In general cohomology and homology are different and are related via the universal coefficient theorem [**37**].

From the perspective of topology, the need to restrict to fields of zero characteristic like $\mathbb{R}$ and $\mathbb{C}$ is a big shortcoming. For example, one would no longer be able to 'detect torsion' and thereby perform basic topological tasks like distinguishing between a torus and a Klein bottle, which is a standard utility of cohomology groups over rings or fields of positive characteristics.

Another technical reason against the Hodge-theoretic approach in Sections 2.2 and 2.3 is that it is not *functorial*. However, if one is primarily interested in engineering and scientific applications, then it is our belief that our approach in Sections 2, 3, and 4 is adequate. Despite our restriction to clique complexes of graphs, our discussions in Section 4 apply verbatim to any simplicial complex.

We should add that although we did not discuss it, one classical use of cohomology and Hodge theory is to deduce topological information about an underlying topological space. Even over a field of characteristic zero, if we sample sufficiently many points V from a sufficiently nice metric space Ω, and set $G = (V, E)$ to be an

appropriately chosen nearest neighbor graph, then

$$\beta_k(G) = \dim H^k(G) = \dim \ker(\Delta_k)$$

gives the number of 'k-dimensional holes' in Ω. While the kernel or 0-eigenspace captures qualitative topological information, the nonzero eigenspaces often capture quantitative geometric information. In the context of graphs [**21, 63**], this is best seen in Δ_0 — its 0-eigenpair tells us whether a graph is connected ($\beta_0(G)$ gives the number of connected components of G, as we saw in Section 3.4) while its smallest nonzero eigenpair tells us how connected the graph is (eigenvalue by way of the Cheeger inequality and eigenvector by way of the Fiedler vector).

5.2. More linear algebra. Since V is a finite set, $L^2(V)$, $L^2_\wedge(E)$, $L^2_\wedge(T)$ are finite-dimensional vector spaces. We may choose bases on these spaces and get

$$L^2(V) \cong \mathbb{R}^p, \quad L^2_\wedge(E) \cong \mathbb{R}^n, \quad L^2_\wedge(T) \cong \mathbb{R}^m$$

where p, n, m are respectively the number of vertices, edges, and triangles in G. Nevertheless, a more natural way would be to regard $L^2_\wedge(E)$ as an n-dimensional subspace of skew-symmetric matrices

$$\big\{[x_{ij}]_{i,j=1}^p \in \mathbb{R}^{p \times p} : x_{ij} = -x_{ji}, \; x_{ij} = 0 \text{ whenever } \{i,j\} \notin E\big\},$$

and $L^2_\wedge(T)$ as an m-dimensional subspace of skew-symmetric hypermatrices [**48**]

$$\big\{[\varphi_{ijk}]_{i,j,k=1}^p \in \mathbb{R}^{p \times p \times p} : \varphi_{ijk} = \varphi_{jki} = \varphi_{kij} = -\varphi_{jik} = -\varphi_{ikj} = -\varphi_{kji},$$
$$\varphi_{ijk} = 0 \text{ whenever } \{i,j,k\} \notin T\big\}.$$

See also Section C for examples of how one may in practice write down matrices representing k-coboundary operators and Hodge k-Laplacians for $k = 0, 1, 2$.

5.3. Continuous analogue. The continuous analogue of Section 3 where the graph G is replaced by a region Ω in $\mathbb{R}^n$ could be illuminating for readers who are familiar with partial differential equations and physics. This section requires nothing more than multivariate calculus.

The Laplace or homogeneous Poisson equation is the PDE

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} = 0$$

over a domain $\Omega \subseteq \mathbb{R}^3$, or more generally,

$$(5.1) \qquad -\Delta_0 f = -\nabla^2 f = -\nabla \cdot \nabla f = \operatorname{div} \operatorname{grad} f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2} = 0$$

over a domain $\Omega \subseteq \mathbb{R}^n$.

The Laplace equation is ubiquitous in physical phenomena, we find it in electrostatics, describing electric potential in free space with no charge; in fluid mechanics, describing the velocity potential of an incompressible fluid; in thermal conduction, as the stationary heat equation without a heat source. Furthermore, it tells us a great deal about the topology and geometry of the domain Ω [**44**].

To ensure uniqueness of solution, one would usually study the Laplace equation with some boundary conditions, say,

$$\begin{cases} \Delta_0 f = 0 & \text{in } \Omega, \\ f = g & \text{on } \partial\Omega, \end{cases} \quad \text{or} \quad \begin{cases} \Delta_0 f = 0 & \text{in } \Omega, \\ \partial f/\partial\nu = g & \text{on } \partial\Omega. \end{cases}$$

0-cohomolgy may be viewed as the study of solutions to Laplace equation with no boundary conditions. A 0-cohomology class is a harmonic function, i.e., a scalar field $f : \Omega \to \mathbb{R}$ that satisfies (5.1). The 0th cohomology group is then the set of all harmonic functions or harmonic 0-forms,

$$H^0(\Omega) = \ker(\Delta_0) = \{f : \Omega \to \mathbb{R} : \Delta_0 f = 0\}.$$

There is also the vector Laplace or homogeneous vector Poisson equation in $\mathbb{R}^3$, usually stated with some boundary conditions:

$$\begin{cases} -\operatorname{grad}\operatorname{div} f + \operatorname{curl}\operatorname{curl} f = 0 & \text{in } \Omega, \\ f \cdot \nu = 0, \quad \operatorname{curl} f \times \nu = 0 & \text{on } \partial\Omega. \end{cases}$$

If we drop the boundary conditions, we get

$$(5.2) \qquad \Delta_1 f = \nabla(\nabla \cdot f) - \nabla \times (\nabla \times f) = \operatorname{curl}\operatorname{curl} f - \operatorname{grad}\operatorname{div} f = 0,$$

where Δ_1 is the *Helmholtz operator* or *vector Laplacian* in $\mathbb{R}^3$.

1-cohomolgy may be viewed as the study of solutions to the vector Laplace equation with no boundary conditions. A 1-cohomology class is a harmonic 1-form, i.e., a vector field $f : \Omega \to \mathbb{R}^3$ that satisfies (5.2). The 1st cohomology group is then the set of all harmonic 1-forms,

$$H^1(\Omega) = \ker(\Delta_1) = \{f : \Omega \to \mathbb{R}^3 : \Delta_1 f = 0\}.$$

5.4. Computations. One particular nice feature of the Hodge decomposition is that it can be efficiently computed by solving least squares problems. For example, to compute the decomposition in (3.7) for any given $X \in L^2_\wedge(E)$, we may solve the two least squares problems

$$\min_{f \in L^2(V)} \|\operatorname{grad} f - X\| \qquad \text{and} \qquad \min_{\Phi \in L^2_\wedge(T)} \|\operatorname{curl}^* \Phi - X\|,$$

to get X_H as $X - \operatorname{grad} f - \operatorname{curl}^* \Phi$. Alternatively, we may solve

$$\min_{Y \in L^2_\wedge(E)} \|\Delta_1 Y - X\|$$

for the minimizer Y and get X_H as the residual $X - \Delta_1 Y$ directly. Once we obtain $\Delta_1 Y$, we may use the decomposition (2.9),

$$\operatorname{im}(\Delta_1) = \operatorname{im}(\operatorname{grad}) \oplus \operatorname{im}(\operatorname{curl}^*),$$
$$\Delta_1 Y = \operatorname{grad} f + \operatorname{curl}^* \Phi,$$

and solve either

$$\min_{f \in L^2(V)} \|\operatorname{grad} f - \Delta_1 Y\| \qquad \text{or} \qquad \min_{\Phi \in L^2_\wedge(T)} \|\operatorname{curl}^* \Phi - \Delta_1 Y\|$$

to get the remaining two components.

We have the choice of practical, efficient, and stable methods like Krylov subspace methods for singular symmetric least squares problems [**19, 20**] or specialized methods for the Hodge 1-Laplacian with proven complexity bounds [**23**].

6. Applications

Applications of cohomology and Hodge theory are plentiful, we find them in numerical analysis [**3**], peridynamics [**30**], topological data analysis [**26**], computational topology [**33**], graphics [**66**], image processing [**70**], robotics [**45**], sensor networks [**65**], neuroscience [**53**], and many other areas in physical science and engineering. But these applications are not 'surprising' in the sense that they all concern physics, geometry, or topology — areas that gave birth to cohomology and Hodge theory in the first place.

What we find somewhat unexpected are recent applications of cohomology and Hodge theory to game theory [**15**] and ranking [**43**]. It is not at first obvious how cohomology and Hodge theory can play a role in these areas, which are less structured compared to applications arising from manifestation of physical laws (e.g. Maxwell's equations or Navier–Stokes equations) naturally amenable to cohomology and Hodge theory. We will briefly describe these in the next two sections.

The headings *game theory* and *ranking* provide basic frameworks that underlie other data analytic or information theoretic applications such as ad auction, social networks (for the former), web search, recommendation systems, crowd sourcing (for the latter). Other recent works along these lines may be found in [**16, 28, 40, 56, 69**]. Cohomology and Hodge theory are more widely applicable than they seem at first.

6.1. Rankings. The application to ranking in [**43**] involves nothing more than attaching ranking-theoretic meanings to the mathematical objects in Section 3. There is a very readable account of this in [**4**].

Suppose V is a set of alternatives to be ranked and $E \subseteq \binom{V}{2}$ is the set of pairs of alternatives that have been compared. The degree of preference of an alternative i over another alternative j is then naturally quantified as $X(i,j)$, since if $X(i,j)$ is the amount one favors i over j, then it is natural to assume that $X(j,i) = -X(i,j)$ is the amount one disfavors j over i. This is the only assumption required of the method.

We call the resulting edge flow $X \in L^2_\wedge(E)$ a *pairwise ranking*. If there is more than one voter, X may be computed by aggregating over all voters using a variety of rank aggregation methods that are standard in Statistics [**24**]. The Helmholtz decomposition (2.6) gives

(6.1)
$$L^2_\wedge(E) = \mathrm{im}(\mathrm{grad}) \oplus \ker(\Delta_1) \oplus \mathrm{im}(\mathrm{curl}^*),$$
$$X \;\;=\;\; \mathrm{grad}\, f \;\;+\;\; X_H \;\;+\;\; \mathrm{curl}^* \Phi,$$

may then be interpreted as

  pairwise ranking = consistent + globally inconsistent + locally inconsistent.

Inconsistency refers to situations where we have circular preference relations like $i \succ j \succ k \succ i$. Consistency means an absence of inconsistencies.

The consistent component $X_P = -\mathrm{grad}\, f$ yields a score[7] function $f \in L^2(V)$ that allows alternatives to be globally ranked: $i \succ j$ iff $f(i) > f(j)$. The total size of the two inconsistent components (measured by their norms) provides a 'certificate of reliability' for the global ranking obtained — if the sum of the inconsistent components is large relative to the consistent component, then the global ranking

---

[7]A score function is a potential function, cf. (3.5), in the context of ranking.

obtained is unreliable. This method of obtaining a global ranking of alternatives via the Helmholtz decomposition is called *HodgeRank*.

The further splitting of the inconsistent component into local and global parts gives finer details about the nature of the inconsistencies. Roughly speaking, the local inconsistent component $\operatorname{curl}\Phi$ measures inconsistencies among items ranked closed together whereas the global inconsistent component X_H measures inconsistencies among items ranked far apart. One may view the two inconsistent components as a quantification of various conundrums in voting theory: Condorcet paradox, Arrow's impossibility, McKelvey chaos, etc [**58**].

In this context, other quantities that we discussed in Section 3 also have ranking-theoretic meanings. For example, assuming we use the standard l^2-inner products, the divergence of a pairwise ranking $X \in L^2_\wedge(E)$ is a function $\operatorname{div} X \in L^2(V)$ and for each alternative $i \in V$,

$$(\operatorname{div} X)(i) = \sum\nolimits_{j=1}^{n} X(i,j)$$

gives a generalization of the *Borda count* [**25, 58**]. In particular, $\operatorname{div} X(i) = 0$ implies that the alternative i is preference-neutral.

Table 4 provides an illustration, where a variant of HodgeRank [**35**] is applied to the dataset from the 2006 Netflix Prize competition [**10**] to obtain a global ranking of the 17,770 movies[8] (the alternatives) based on 100,480,507 ratings given by 480,189 viewers (the voters). The middle and right columns show the top 15 results obtained when HodgeRank is used in conjunction with the log-odds model and linear model for rank-aggregation. The left column shows the ranking obtained via simple averaging of the ratings; the shortcoming of such an approach is obvious — if Movie X gets one single 5-star rating and Movie Y gets 10,000 5-star ratings and one single 1-star rating, ranking by average ratings would put Movie X ahead of Movie Y. We call such ratings *imbalanced*. The results shown here already includes a mitigating factor whereby infrequently rated movies are filtered out beforehand but the imbalanced effect is still visible in the left column. One reason the effect is not as pronounced in HodgeRank is that it uses an inner product with

$$w_{ij} = w_{ji} = \text{number of voters who have compared alternatives } i \text{ and } j$$

for the weights in (3.1) to correct for such imbalanced ratings.

6.2. Games. Consider a game with n players, indexed $1, \ldots, n$. We will define a graph $G = (V, E)$ as follows. The vertex set V is a finite set of the form

$$V = S_1 \times \cdots \times S_n$$

where S_i is a finite set of all possible *strategies* for the ith player. An element $(s_1, \ldots, s_n) \in V$ is called a *strategy profile* where $s_i \in S_i$ denotes the strategy of the ith player. Once we have specified V, the edge set $E \subseteq \binom{V}{2}$ is completely determined in the following way: for any pair of vertices $(s_1, \ldots, s_n), (t_1, \ldots, t_n) \in V$,

(6.2) $\{(s_1, \ldots, s_n), (t_1, \ldots, t_n)\} \in E$ iff there exists exactly one i where $s_i \neq t_i$.

In other words, E comprises all pairs of strategy profiles that differ only in the strategy of a single player. Note that in this application E cannot be arbitrary but is completely determined by V.

---

[8]LOTR III appears twice because of the two DVD editions, theatrical and extended.

TABLE 4. Top 15 movies in the Netflix dataset generated by a variant of HodgeRank (middle and right). For comparison, left column is a ranking via the mean ratings of the movies. (Reproduced verbatim from [**35**, Table 2, p. 62])

MEAN RATINGS	HODGERANK/LOG-ODDS	HODGERANK/LINEAR
LOTR III: Return ...	LOTR III: Return ...	LOTR III: Return ...
LOTR I: The Fellowship ...	LOTR I: The Fellowship ...	LOTR I: The Fellowship ...
LOTR II: The Two ...	LOTR II: The Two ...	LOTR II: The Two ...
Lost: Season 1	Star Wars V: Empire ...	Lost: S1
Battlestar Galactica: S1	Raiders of the Lost Ark	Star Wars V: Empire ...
Fullmetal Alchemist	Star Wars IV: A New Hope	Battlestar Galactica: S1
Trailer Park Boys: S4	Shawshank Redemption	Star Wars IV: A New Hope
Trailer Park Boys: S3	Star Wars VI: Return ...	LOTR III: Return ...
Tenchi Muyo! ...	LOTR III: Return ...	Raiders of the Lost Ark
Shawshank Redemption	The Godfather	The Godfather
Veronica Mars: S1	Toy Story	Shawshank Redemption
Ghost in the Shell: S2	Lost: S1	Star Wars VI: Return ...
Arrested Development: S2	Schindler's List	Gladiator
Simpsons: S6	Finding Nemo	Simpsons: S5
Inu-Yasha	CSI: S4	Schindler's List

Once we have specified V and therefore E, a game on the graph $G = (V, E)$ is determined by the *utility function* $f_i \in L^2(V)$ of the ith player, $i = 1, \ldots, n$. We write $F := (f_1, \ldots, f_n)$. So $F : V \to \mathbb{R}^n$ is a vector field on V and we will call it the *utility vector field* of the game. We denote the set of all vector fields by

$$L^2(V; \mathbb{R}^n) := \{F : V \to \mathbb{R}^n : f_1, \ldots, f_n \in L^2(V)\} \cong L^2(V) \oplus \cdots \oplus L^2(V).$$

Since every game on G is specified by a utility vector field, $L^2(V; \mathbb{R}^n)$ parameterizes all games on G.

A *potential game* is one where the utility vector field $F = (f_1, \ldots, f_n)$ satisfies

$$\operatorname{grad} f_1 = \cdots = \operatorname{grad} f_n.$$

This condition implies that f_i and f_j differ at most by an additive constant[9]. It may be interpreted to mean that the utilities of all players are aligned with a global objective. Such games were first proposed in [**54**] and are widely studied in game theory. They are easy to analyze and have *pure Nash equilibria*, i.e., strategy profiles $(s_1, \ldots, s_n) \in V$ such that

$$f_i(s_1, \ldots, s_i, \ldots, s_n) \geq f_i(s_1, \ldots, s_i', \ldots, s_n)$$

for every $s_i' \in S_i$ and for each $i = 1, \ldots, n$.

A *harmonic game* is one where the utility vector field $F = (f_1, \ldots, f_n)$ satisfies

$$\Delta_0(f_1 + \cdots + f_n) = 0,$$

i.e., $f_1 + \cdots + f_n \in \ker(\Delta_0)$ is a harmonic function. These were first defined in [**15**] based on similar earlier ideas [**42**]. Such games generically have no pure Nash

---

[9]By the definition of E, G is connected and so $\operatorname{grad} f = 0$ implies f is a constant function.

equilibrium and may be regarded essentially as sums of cycles [**15**]. A well-known simple example would be the rock-paper-scissors game.

We define subspaces of utility vector fields corresponding to potential and harmonic games respectively:

$$P = \{F \in L^2(V; \mathbb{R}^n) : \operatorname{grad} f_i = \operatorname{grad} f_j \text{ for all } i, j = 1, \ldots, n\},$$
$$H = \{F \in L^2(V; \mathbb{R}^n) : \Delta_0(f_1 + \cdots + f_n) = 0\},$$
$$C = \{F \in L^2(V; \mathbb{R}^n) : f_1, \ldots, f_n \text{ constant functions}\}.$$

As before, $F = (f_1, \ldots, f_n)$. C is the set of constant vector fields and it is easy to see that $P \cap H = C$.

The *Jacobian* is $\operatorname{Jac} : L^2(V; \mathbb{R}^n) \to L^2_\wedge(E; \mathbb{R}^n)$ takes $F \in L^2(V; \mathbb{R}^n)$ to

$$\operatorname{Jac} F = (\operatorname{grad} f_1, \ldots, \operatorname{grad} f_n),$$

in $L^2_\wedge(E; \mathbb{R}^n) := \{(X_1, \ldots, X_n) : X_1, \ldots, X_n \in L^2_\wedge(E)\}$. In the context of games, an edge flow in $L^2_\wedge(E)$ is called a *game flow* and an element of $L^2_\wedge(E; \mathbb{R}^n)$ is a vector field of game flows. The image of the spaces P, H, C under the Jacobian are respectively,

$$\operatorname{Jac}(P) = \{(X, \ldots, X) \in L^2_\wedge(E; \mathbb{R}^n) : X \in L^2_\wedge(E)\},$$
$$\operatorname{Jac}(H) = \{(X_1, \ldots, X_n) \in L^2_\wedge(E; \mathbb{R}^n) : \operatorname{div}(X_1 + \cdots + X_n) = 0\},$$

and $\operatorname{Jac}(C) = \{0\}$.

In [**15**], the reader would find a decomposition of games into potential, harmonic, and nonstrategic components, induced by the Helmholtz decomposition of *game flows*. The details are a bit too involved for this article and we would instead just describe the latter. The main distinction is that a decomposition of games is a decomposition of $L^2(V; \mathbb{R}^n)$ whereas a decomposition of game flows is a decomposition of $L^2_\wedge(E)$.

A departure from the general Helmholtz decomposition is that game flows are always curl-free, i.e., always in ker(curl). So the Helmholtz decomposition of game flows takes the form

$$\ker(\operatorname{curl}) = \operatorname{im}(\operatorname{grad}) \oplus \ker(\Delta_1),$$
$$X \quad = \quad X_P \quad + \quad X_H,$$

with the following interpretation,

$$\text{game flow} = \text{potential game flow} + \text{harmonic game flow}.$$

EXAMPLE 6.1 (Road sharing game). We discuss a beautiful three-player game with nonzero potential and harmonic components first given in [**15**]. In this game, three players C, R, P each chooses one of two roads a, b.

The scenario is the immediate aftermath of a bank robbery where a *policeman* (player P) is in pursuit of the bank *robber* (player R). Naturally the policeman wants to be on the same road as the robber while the robber wants to avoid that. The *commuter* (player C), an innocent bystander, does not want collateral damage by being on the same road as either the robber or the policeman.

The respective utility functions, or rather, changes to the values of the utility functions (i.e., payoffs), are described as follows:

- commuter's payoff decreases by 2 with each other player sharing his road;
- robber's payoff is -1 if the policeman shares his road and 0 otherwise;
- policeman's payoff is exactly the negative of the robber's payoff.

Here $S_C = S_R = S_P = \{a, b\}$ and the set of strategy profiles is therefore

$$V = \{(s_C, s_R, s_P) : s_C \in S_C, \ s_R \in S_R, \ s_P \in S_P\}$$
$$= \{(a,a,a),(a,a,b),(a,b,a),(a,b,b),(b,a,a),(b,a,b),(b,b,a),(b,b,b)\}.$$

E is uniquely determined by V according to (6.2) and comprises pairs of strategy profiles that differ only in the strategy of a single player:

$$\{(a,a,a),(a,a,b)\}, \ \{(a,a,a),(a,b,a)\}, \ \{(a,a,a),(b,a,a)\}, \ \{(a,a,b),(a,b,b)\},$$
$$\{(a,a,b),(b,a,b)\}, \ \{(a,b,a),(a,b,b)\}, \ \{(a,b,a),(b,b,a)\}, \ \{(b,a,a),(b,a,b)\},$$
$$\{(b,a,a),(b,b,a)\}, \ \{(a,b,b),(b,b,b)\}, \ \{(b,a,b),(b,b,b)\}, \ \{(b,b,a),(b,b,b)\}.$$
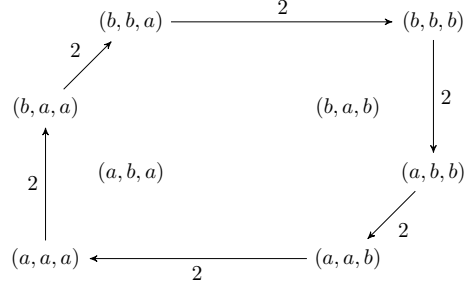
The Helmholtz decomposition of this game is shown in Figure 8. An edge flow between two vertices (i.e., strategy profiles) is represented by an arrow and a weight — the direction of the arrow represents an improvement in the payoff of the player who changes its strategy; the amount of improvement is given by the corresponding weight.
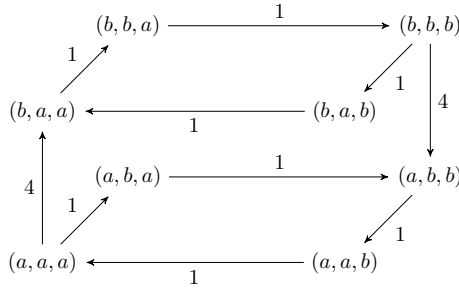
(A) flow of road-sharing game
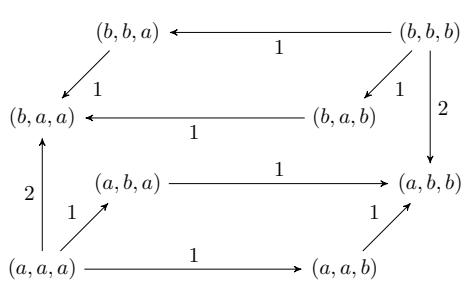
(B) potential component

(C) harmonic component

FIGURE 8. Helmholtz decomposition of the road sharing game. (Reproduced with slight modification from [**15**, Figure 1, p. 475])

Observe that the game flow $X \in L^2_\wedge(E)$ shown in Figure 8A is indeed curl-free since $T = \varnothing$ and so $L^2_\wedge(T) = \{0\}$. The potential component X_P in Figure 8B flows out of sources (a,a,a) and (b,b,b) into sinks (a,b,b) and (b,a,a). An explicit potential function $f \in L^2(V)$ is easy to construct, for example,

$$f(a,a,a) = f(b,b,b) = 1, \quad f(a,b,b) = f(b,a,a) = -1,$$
$$f(a,a,b) = f(a,b,a) = f(b,a,b) = f(b,b,a) = 0.$$

Check that $X_P = -\operatorname{grad} f$, as in (3.5).

The harmonic component $X_H \in L^2_\wedge(E)$ shown in Figure 8C is evidently a cycle[10] that goes from

$$(a,a,a) \to (b,a,a) \to (b,b,a) \to (b,b,b) \to (a,b,b) \to (a,a,b) \to (a,a,a)$$

in a loop. The divergence of X_H is easily seen to be zero by (3.4) since the inflow equals the outflow at every vertex. Given that X and therefore X_H is in $\ker(\operatorname{curl})$, we conclude that $X_H \in \ker(\operatorname{curl}) \cap \ker(\operatorname{div}) = \ker(\Delta_1)$ by (3.7), i.e., it is indeed a harmonic 1-cochain.

7. Nonlinear Laplacians on a graph

In this last section, we will introduce another generalization of the graph Laplacian that is also constructed out of the coboundary operators in Section 3.3. Whereas the Hodge Laplacians may be viewed as a *higher order* generalization of the graph Laplacian, acting linearly on functions of higher order cliques as opposed to functions of vertices (i.e., 1-cliques), the generalization in this section takes things in a different direction — they act on functions of vertices but are *nonlinear* operators.

The nonlinear p-Laplacians are defined on open subsets of $\mathbb{R}^n$ [**32, 49**] by

$$(7.1) \qquad L_p f = -\operatorname{div}(\|\operatorname{grad} f\|^{p-2} \operatorname{grad} f)$$

for $1 < p < \infty$, and

$$(7.2) \qquad L_1 f = -\operatorname{div}\left(\frac{\operatorname{grad} f}{\|\operatorname{grad} f\|}\right), \qquad L_\infty f = -(\operatorname{grad} f)^\mathsf{T}(\operatorname{grad}\|\operatorname{grad} f\|^2)$$

for $p = 1, \infty$ respectively. Here $\operatorname{grad} f = \nabla f$ is taken to be a column vector and $\|\cdot\|$ refers to the usual Euclidean norm on $\mathbb{R}^n$, i.e.,

$$\|\operatorname{grad} f\|^2 = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\right)^2.$$

These are all nonlinear operators with the exception of the case $p = 2$, which is a linear operator. The nonlinear 1-Laplacian is sometimes also called the *mean curvature operator*.

For $1 \le p \le \infty$, the *nonlinear p-Laplacians* on a graph $G = (V, E)$ are defined by the same formulas in (7.1) and (7.2) except that the gradient and divergence operators div and grad should now be taken to be the coboundary operators introduced in Section 3.3. In particular, the (non)linear 2-Laplacian is exactly the Hodge 0-Laplacian defined in Section 4.3,

$$L_2 = -\operatorname{div}\operatorname{grad} = \Delta_0,$$

which is of course the usual graph Laplacian. Note that as before, our graphs are undirected and have no loops or multiple edges

For $1 \le p < 2$, L_p has a singularity whenever $\operatorname{grad} f$ vanishes. To define it everywhere we borrow some ideas from nondifferentiable optimization [**22**]. For $p = 1$, the nonlinear 1-Laplacian [**38**] on a graph is given by

$$L_1 f = -\operatorname{div}(\operatorname{sgn}(\operatorname{grad} f)),$$

---

[10]Or more accurately, a cocycle, since we are talking about functions. The directed edges form a 1-cycle but functions on them form 1-cocycles. See Tables 1 and 3.

essentially obtained from (7.2) by replacing $x/\|x\|$ with $\mathrm{sgn}(x)$. The function sgn is a set-valued function defined on $\mathbb{R}$ by

$$\mathrm{sgn}(x) = \begin{cases} -1 & \text{if } x < 0, \\ [-1, 1] & \text{if } x = 0, \\ +1 & \text{if } x > 0, \end{cases}$$

and extended to $\mathbb{R}^n$ coordinatewise, i.e.,

$$\mathrm{sgn}(x) = (\mathrm{sgn}(x_1), \ldots, \mathrm{sgn}(x_n))$$

for $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$. This is the subdifferential of the 1-norm on $\mathbb{R}^n$, i.e., $\partial\|x\|_1 = \mathrm{sgn}(x)$.

An example of an instance where the nonlinear 1-Laplacian appears is the following. In spectral graph theory, a well-known fundamental result is that the isoperimetric or Cheeger constant,

$$h(G) = \min_{S \subset V} \frac{|E(S, V \setminus S)|}{\min(\mathrm{Vol}(S), \mathrm{Vol}(V \setminus S))},$$

of a connected graph G satisfies [**2**, **47**] the Cheeger inequality

$$\frac{1}{2}\lambda_2(L_2) \le h(G) \le \sqrt{2\lambda_2(L_2)},$$

where $\lambda_2(L_2)$ is the second eigenvalue of the usual graph Laplacian. But what if we want equality? The answer, as demonstrated in [**38**, **18**], is that if we replace L_2 by L_1, then we get

$$\lambda_2(L_1) = h(G).$$

We will end this section with a few words about applications. The continuous nonlinear 1-Laplacian, usually under its better-know name 'mean curvature operator,' plays a central role in the Level Set Method [**55**]. In a similar spirit, the discrete version has also been found [**46**] to have a game-theoretic interpretation — the so-called balancing games in $\mathbb{R}^n$ [**62**] may be regarded as a discrete motion by mean curvature. A simple example of a balancing game is as follows: Player I chooses a direction, i.e., a unit vector $x_i \in \mathbb{R}^n$, and Player II chooses whether to flip it, i.e., $\varepsilon_i \in \{-1, +1\}$, in a way so that $\left\|\sum_{i=1}^{t} \varepsilon_i x_i\right\|$ is minimized at step t.

8. Conclusion

Traditional applied mathematics largely involves using partial differential equations to model physical phenomena and traditional computational mathematics largely revolves around numerical solutions of PDEs.

However, one usually needs substantial and rather precise knowledge about a phenomenon in order to write it down as PDEs. For example, one may need to know the dynamical laws (e.g. laws of motions, principle of least action, laws of thermodynamics, quantum mechanical postulates, etc) or conservation laws (e.g. of energy, momentum, mass, charge, etc) underlying the phenomenon before being able to 'write down' the corresponding PDEs (as equations of motion, of continuity and transfer, constitutive equations, field equations, etc). In traditional applied mathematics, it is often taken for granted that there are known physical laws behind the phenomena being modeled.

In modern data applications, this is often a luxury. For example, if we want to build a spam filter, then it is conceivable that we would want to understand the 'laws

of emails.' But we would quickly come to the realization that these 'laws of emails' would be too numerous to enumerate and too inexact to be formulated precisely, even if we restrict ourselves to those relevant for identifying spam. This is invariably the case for any human generated data: movie ratings, restaurant reviews, browsing behavior, clickthrough rates, newsfeeds, tweets, blogs, instagrams, status updates on various social media, etc, but the perhaps surprising thing is that it also applies to many data sets from modern biology and medicine.

For such data sets, all one has is often a rough measure of how similar two data points are and how the dataset is distributed. Topology can be a useful tool in such contexts [**17**] since it requires very little — essentially just a weak notion of separation, i.e., is there a non-trivial open set that contains those points?

If the data set is discrete and finite, which is almost always the case in applications, we can even limit ourselves to simplicial topology, where the topological spaces are simplicial complexes (cf. Section 3.1). Without too much loss of generality, these may be regarded as clique complexes of graphs (cf. Section 5.1): data points are vertices in V and proximity is characterized by cliques: a pair of data points are near each other iff they form an edge in E, a triplet of data points are near one another iff they form a triangle in T, and so on.

In this article, we have undertaken the point of view that

$$\boxed{\text{graphs are discrete Riemannian manifolds}}$$

and that cohomology is a much-relaxed discrete analogue of PDEs (cf. Sections 3.3, 5.3). Furthermore, standard partial differential operators on Riemannian manifolds — gradient, divergence, curl, Jacobian, Hessian, Laplace and Helmholtz operators, Hodge and nonlinear Laplacians — all have natural counterparts on graphs (cf. Sections 3, 4, 7) with useful roles in data applications (cf. Section 6). An example of a line of work that carries this point of view to great fruition is [**8, 9, 52**]. Also, in this article we have only scratched the surface of cohomological and Hodge theoretic techniques in graph theory; see [**57**] for results that go much further.

In traditional computational mathematics, discrete PDEs arise as discretization of continuous PDEs, intermediate by-products of numerical schemes. But in data analytic applications, discrete PDEs tend to play a more central and direct role. Despite these differences, the discrete partial differential operators on graphs introduced in this article may perhaps serve as a bridge on which insights from traditional applied and computational mathematics could cross over and be brought to bear on modern data analytic applications.

Acknowledgments

References

[1] *Manual for Short Course Organizers and Lecturers*, AMS, Providence, RI, 2014, `http://www.ams.org/meetings/short-courses/2014_short_course_manual_final.pdf`.

[2] N. Alon and V. D. Milman, "λ_1, isoperimetric inequalities for graphs, and superconcentrators," *J. Combin. Theory*, Ser. B, **38** (1985), no. 1, pp. 73–88.

[3] D. Arnold, R. S. Falk, and R. Winther, "Finite element exterior calculus: from Hodge theory to numerical stability," *Bull. Amer. Math. Soc.*, **47** (2010), no. 2, pp. 281–354.

[4] D. Austin, "Who's number 1? Hodge theory will tell us," *AMS Feature Column*, December 2012, `http://www.ams.org/samplings/feature-column/fc-2012-12`.

[5] L. Babai, "Automorphism groups, isomorphism, reconstruction," Chapter 27, pp. 1447–1540, in R. L. Graham, M. Grötschel, and L. Lovász (Eds), *Handbook of Combinatorics*, **2**, Elsevier, Amsterdam, Netherlands, 1995.

[6] L. Babai, "Graph isomorphism in quasipolynomial time," *Combinatorics and Theoretical Computer Science Seminar*, University of Chicago, Chicago, IL, November 10, 2015. `http://people.cs.uchicago.edu/~laci/quasipoly.html`

[7] L. Bartholdi, T. Schick, N. Smale, and S. Smale, "Hodge theory on metric spaces," *Found. Comput. Math.*, **12** (2012), no. 1, pp. 1–48.

[8] I. Benjamini and L. Lovász, "Harmonic and analytic functions on graphs," *J. Geom.*, **76** (2003), no. 1–2, pp. 3–15.

[9] I. Benjamini and O. Schramm, "Harmonic functions on planar and almost planar graphs and manifolds, via circle packings," *Invent. Math.*, **126** (1996), no. 3, pp. 565–587.

[10] J. Bennett and S. Lanning, "The Netflix prize," *Proc. KDD Cup Workshop*, pp. 3–6, ACM, New York, NY, 2007.

[11] J. Bentley, "Bumper-sticker computer science," *Commun. ACM*, **28** (1985), no. 9, pp. 896–901.

[12] A. Bossavit, "Whitney forms: a class of finite elements for three-dimensional computations in electromagnetism," *Proc. IEEE*, **135** (1988), no. 8, pp. 493–500.

[13] L. M. Butler, "A beautiful mind," *Notices Amer. Math. Soc.*, **49** (2002), no. 4, pp. 455–457.

[14] S. Butler and J. Grout, "A construction of cospectral graphs for the normalized Laplacian," *Electron. J. Combin.*, **18** (2011), no. 1, Paper 231, 20 pp.

[15] O. Candogan, I. Menache, A. Ozdaglar, and P. Parrilo, "Flows and decompositions of games: harmonic and potential games," *Math. Oper. Res.*, **36** (2011), no. 3, pp. 474–503.

[16] O. Candogan, A. Ozdaglar, and P. Parrilo, "Dynamics in near-potential games," *Games Econom. Behav.*, **82** (2013), pp. 66–90.

[17] G. Carlsson, "Topological pattern recognition for point cloud data," *Acta Numer.*, **23** (2014), pp. 289–368.

[18] K. C. Chang, "Spectrum of the 1-Laplacian and Cheeger's constant on graphs," *J. Graph Theory*, (2015), to appear.

[19] S.-C. T. Choi and M. A. Saunders, "Algorithm 937: MINRES-QLP for symmetric and Hermitian linear equations and least-squares problems," *ACM Trans. Math. Software*, **40** (2014), no. 2, Art. 16, 12 pp.

[20] S.-C. T. Choi, C. C. Paige, and M. A. Saunders, "MINRES-QLP: a Krylov subspace method for indefinite or singular symmetric systems," *SIAM J. Sci. Comput.*, **33** (2011), no. 4, pp. 1810–1836.

[21] F. R. K. Chung, *Spectral Graph Theory*, CBMS Regional Conference Series in Mathematics, **92**, AMS, Providence, RI, 1997.

[22] F. H. Clarke, *Optimization and Nonsmooth Analysis*, 2nd Ed., Classics in Applied Mathematics, **5**, SIAM, Philadelphia, PA, 1990.

[23] M. B. Cohen, B. T. Fasy, G. L. Miller, A. Nayyeri, R. Peng, and N. Walkington, "Solving 1-Laplacians of convex simplicial complexes in nearly linear time," *Proc. Annual ACM–SIAM Symp. Discrete Algorithms* (SODA '14), **25** (2014), pp. 204–216.

[24] H. A. David, *The Method of Paired Comparisons*, 2nd Ed., Griffin's Statistical Monographs and Courses, **41**, Oxford University Press, New York, NY, 1988.

[25] J.-C. De Borda, "Mémoire sur les élections au scrutin," *Histoire de'Académie Royale des Sciences*, **102** (1781), pp. 657–665.

[26] V. De Silva, D. Morozov, and M. Vejdemo-Johansson, "Persistent cohomology and circular coordinates," *Discrete Comput. Geom.*, **45** (2011), no. 4, pp. 737–759.

[27] M. Desbrun, A. Hirani, M. Leok, and J. E. Marsden, "Discrete exterior calculus," `http://arxiv.org/abs/math/0508341`, *preprint*, (2005), 53 pp.

[28] T. Dey, A. Hirani, and B. Krishnamoorthy, "Optimal homologous cycles, total unimodularity, and linear programming," *SIAM J. Comput.*, **40** (2011), no. 4, pp. 1026–1044.

[29] J. Dodziuk, "Finite-difference approach to the Hodge theory of harmonic forms," *Amer. J. Math.*, **98** (1976), no. 1, pp. 79–104.

[30] Q. Du, M. Gunzburger, R. Lehoucq, and K. Zhou, "A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws," *Math. Models Methods Appl. Sci.*, **23** (2013), no. 3, pp. 493–540.

[31] B. Eckmann, "Harmonische funktionen und randwertaufgaben in einem komplex," *Comment. Math. Helv.*, **17**, (1945), pp. 240–255.

[32] L. C. Evans, *Partial Differential Equations*, 2nd Ed., Graduate Studies in Mathematics, **19**, American Mathematical Society, Providence, RI, 2010.

[33] J. Friedman, "Computing Betti numbers via combinatorial Laplacians," *Algorithmica*, **21** (1998), no. 4, pp. 331–346.

[34] P. E. Gill, W. Murray, and M. H. Wright, *Numerical Linear Algebra and Optimization*, Addison–Wesley, Redwood City, CA, 1991.

[35] D. Gleich and L.-H. Lim, "Rank aggregation via nuclear norm minimization," *Proc. ACM SIGKDD Conf. Knowledge Discovery and Data Mining* (KDD '11), **17** (2011), pp. 60–68.

[36] J.-M. Guo, J. Li, and W. C. Shiu, "On the Laplacian, signless Laplacian and normalized Laplacian characteristic polynomials of a graph," *Czechoslovak Math. J.*, **63** (2013), no. 3, pp. 701–720.

[37] A. Hatcher, *Algebraic Topology*, Cambridge University Press, Cambridge, UK, 2002.

[38] M. Hein and T. Bühler, "An inverse power method for nonlinear eigenproblems with applications in 1-spectral clustering and sparse PCA," *Proc. Adv. Neural Inform. Process. Sys.* (NIPS 2010), **23** (2010), pp. 847–855.

[39] A. Hirani, *Discrete Exterior Calculus*, Ph.D. thesis, California Institute of Technology, Pasadena, CA, 2003.

[40] A. Hirani, K. Kalyanaraman, and S. Watts, "Least squares ranking on graphs," `http://arxiv.org/abs/1011.1716`, *preprint*, (2011), 43 pp.

[41] W. V. D. Hodge, *The Theory and Applications of Harmonic Integrals*, 2d Ed., Cambridge University Press, Cambridge, 1952.

[42] J. Hofbauer and K. H. Schlag, "Sophisticated imitation in cyclic games," *J. Evolutionary Econom.*, **10** (2000), no. 5, pp. 523–543.

[43] X. Jiang, L.-H. Lim, Y. Yao, and Y. Ye, "Statistical ranking and combinatorial Hodge theory," *Math. Program.*, **127** (2011), no. 1, pp. 203–244.

[44] M. Kac, "Can one hear the shape of a drum?" *Amer. Math. Monthly*, **73** (1966), no. 4, pp. 1–23.

[45] P. Kingston and M. Egerstedt, "Distributed-infrastructure multi-robot routing using a Helmholtz–Hodge decomposition," *Proc. IEEE Conf. Decis. Control* (CDC), **50** (2011), pp. 5281–5286.

[46] R. V. Kohn and S. Serfaty, "Second-order PDE's and deterministic games," *Proc. Internat. Congress Ind. Appl. Math.* (ICIAM '07), pp. 239–249, EMS, Zürich, 2009.

[47] G. Lawler and A. D. Sokal, "Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger's inequality," *Trans. Amer. Math. Soc.*, **309** (1988), no. 2, pp. 557–580.

[48] L.-H. Lim, "Tensors and hypermatrices," Art. 15, pp. 15-1–15-30, in L. Hogben (Ed.), *Handbook of Linear Algebra*, 2nd Ed., CRC Press, Boca Raton, FL, 2013.

[49] P. Lindqvist, *Notes on the p-Laplace Equation*, `http://www.math.ntnu.no/~lqvist/p-laplace.pdf`, Trondheim, Norway, 2006.

[50] H. J. Lipkin, *Beta Decay for Pedestrians*, Dover, Mineola, NY, 2004.

[51] H. J. Lipkin, *Lie Groups for Pedestrians*, 2nd Ed., Dover, Mineola, NY, 2002.

[52] L. Lovász, "Discrete analytic functions: an exposition," pp. 241–273, in A. Grigoryan and S. T. Yau (Eds), *Surveys in Differential Geometry: Eigenvalues of Laplacians and Other Geometric Operators*, **IX**, International Press, Somerville, MA, 2004.

[53] W. Ma, J.-M. Morel, S. Osher, and A. Chien, "An L_1-based variational model for retinex theory and its applications to medical images," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* (CVPR), **29** (2011), pp. 153–160.

[54] D. Monderer and L. S. Shapley, "Potential games," *Games Econom. Behav.*, **14** (1996), no. 1, pp. 124–143.

[55] S. Osher and J. A. Sethian, "Fronts propagating with curvature-dependent speed: algorithms based on Hamilton–Jacobi formulations," *J. Comput. Phys.*, **79** (1988), no. 1, pp. 12–49.

[56] B. Osting, J. Darbon, and S. Osher, "Statistical ranking using the ℓ^1-norm on graphs," *Inverse Probl. Imaging*, **7** (2013), no. 3, pp. 907–926.

[57] P. Hanlon, "The Laplacian method," pp. 65–91, in S. Fomin (Ed), *Symmetric Functions 2001: Surveys of Developments and Perspectives*, NATO Science Series II: Mathematics, Physics and Chemistry, **74**, Kluwer, Dordrecht, Netherlands, 2002.

[58] D. Saari, *Geometry of Voting*, Springer-Verlag, Berlin, 1994.

[59] H. M. Schey, *Div, Grad, Curl, and All That: An informal text on vector calculus*, 4th Ed., Norton, New York, NY, 2005.

[60] B. Simon, "OPUC on one foot," *Bull. Amer. Math. Soc.*, **42** (2005), no. 4, pp. 431–460.

[61] N. Smale and S. Smale, "Abstract and classical Hodge-de Rham theory," *Anal. Appl.*, **10** (2012), no. 1, pp. 91–111.

[62] J. Spencer, "Balancing games," *J. Combinatorial Theory*, Ser. B, **23** (1977), no. 1, pp. 68–74.

[63] D. Spielman, "Spectral graph theory," Chapter 18, pp. 495–524, in U. Naumann and O. Schenk (Eds), *Combinatorial Scientific Computing*, CRC Press, Boca Raton, FL, 2012.

[64] G. Strang, "The fundamental theorem of linear algebra," *Amer. Math. Monthly*, **100** (1993), no. 9, pp. 848–855.

[65] A. Tahbaz-Salehi and A. Jadbabaie, "Distributed coverage verification in sensor networks without location information," *IEEE Trans. Autom. Control*, **55** (2010), no. 8, pp. 1837–1849.

[66] Y. Tong, S. Lombeyda, A. Hirani, and M. Desbrun, "Discrete multiscale vector field decomposition," *ACM Trans. Graph.*, **22** (2003), no. 3, pp. 445–452.

[67] F. W. Warner, *Foundations of Differentiable Manifolds and Lie Groups*, Graduate Texts in Mathematics, **94**, Springer-Verlag, New York, NY, 1983.

[68] H. Whitney, *Geometric Integration Theory*, Princeton University Press, Princeton, NJ, 1957.

[69] Q. Xu, Q. Huang, T. Jiang, B. Yan, Y. Yao, and W. Lin, "HodgeRank on random graphs for subjective video quality assessment," *IEEE Trans. Multimedia*, **14** (2012), no. 3, pp. 844–857.

[70] S. Yu, "Angular embedding: A robust quadratic criterion," *IEEE Trans. Pattern Anal. Mach. Intell.*, **34** (2012), no. 1, pp. 158–173.

Appendix A. Numerical Linear Algebra Refresher

By *numerical linear algebra*, we mean linear algebra over any subfield of $\mathbb{C}$. We will provide routine proofs for some numerical linear algebraic facts that we have used freely in Section 2. We will work over $\mathbb{R}$ for convenience.

THEOREM A.1. *Let $A \in \mathbb{R}^{m \times n}$. Then*

① $\ker(A^*A) = \ker(A)$,

② $\operatorname{im}(A^*A) = \operatorname{im}(A^*)$,

③ $\ker(A^*) = \operatorname{im}(A)^\perp$,

④ $\operatorname{im}(A^*) = \ker(A)^\perp$,

⑤ $\mathbb{R}^n = \ker(A) \oplus \operatorname{im}(A^*)$.

PROOF.

① Clearly $\ker(A) \subseteq \ker(A^*A)$. If $A^*Ax = 0$, then $\|Ax\|^2 = x^*A^*Ax = 0$, so $Ax = 0$, and so $\ker(A^*A) \subseteq \ker(A)$.

② Applying rank-nullity theorem twice with ①, we get

$$\operatorname{rank}(A^*A) = n - \operatorname{nullity}(A^*A)$$
$$= n - \operatorname{nullity}(A) = \operatorname{rank}(A) = \operatorname{rank}(A^*).$$

Since $\operatorname{im}(A^*A) \subseteq \operatorname{im}(A^*)$, the result follows.

③ If $x \in \operatorname{im}(A)^\perp$, then $0 = \langle x, Ay \rangle = \langle A^*x, y \rangle$ for all $y \in \mathbb{R}^n$, so $A^*x = 0$. If $x \in \ker(A^*)$, then $\langle x, Ay \rangle = \langle A^*x, y \rangle = 0$ for all $y \in \mathbb{R}^n$, so $x \in \operatorname{im}(A)^\perp$.

④ By ③, $\operatorname{im}(A^*)^\perp = \ker(A^{**}) = \ker(A)$ and result follows.

⑤ $\mathbb{R}^n = \ker(A) \oplus \ker(A)^\perp = \ker(A) \oplus \operatorname{im}(A^*)$ by ④. $\qquad \square$

Our next proof ought to convince readers that the Hodge decomposition theorem is indeed an extension of the Fredholm alternative theorem to a pair of matrices.

THEOREM A.2. *Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ with $AB = 0$. Then*

    ⑥ $\ker(A^*A + BB^*) = \ker(A) \cap \ker(B^*)$,

    ⑦ $\ker(A) = \operatorname{im}(B) \oplus \ker(A^*A + BB^*)$,

    ⑧ $\ker(B^*) = \operatorname{im}(A^*) \oplus \ker(A^*A + BB^*)$,

    ⑨ $\mathbb{R}^n = \operatorname{im}(A^*) \oplus \ker(A^*A + BB^*) \oplus \operatorname{im}(B)$,

    ⑩ $\operatorname{im}(A^*A + BB^*) = \operatorname{im}(A^*) \oplus \operatorname{im}(B)$.

PROOF. Note that $\operatorname{im}(B) \subseteq \ker(A)$ as $AB = 0$, $\operatorname{im}(A^*) \subseteq \ker(B^*)$ as $B^*A^* = 0$.

    ⑥ Clearly $\ker(A) \cap \ker(B^*) \subseteq \ker(A^*A + BB^*)$. Let $x \in \ker(A^*A + BB^*)$. Then $A^*Ax = -BB^*x$.
- Multiplying by A, we get $AA^*Ax = -ABB^*x = 0$ since $AB = 0$. So $A^*Ax \in \ker(A)$. But $A^*Ax \in \operatorname{im}(A^*) = \ker(A)^\perp$ by ④. So $A^*Ax = 0$ and $x \in \ker(A^*A) = \ker(A)$ by ①.
- Multiplying by B^*, we get $0 = B^*A^*Ax = -B^*BB^*x$ since $B^*A^* = 0$. So $BB^*x \in \ker(B^*)$. But $BB^*x \in \operatorname{im}(B) = \ker(B^*)^\perp$ by ③. So $BB^*x = 0$ and $x \in \ker(BB^*) = \ker(B^*)$ by ①.

    Hence $x \in \ker(A) \cap \ker(B^*)$.

    ⑦ Applying ⑤ to B^*,

$$\ker(A) = \mathbb{R}^n \cap \ker(A) = [\ker(B^*) \oplus \operatorname{im}(B)] \cap \ker(A)$$
$$= [\ker(B^*) \cap \ker(A)] \oplus [\operatorname{im}(B) \cap \ker(A)]$$
$$= \ker(A^*A + BB^*) \oplus \operatorname{im}(B),$$

    where the last equality follows from ⑥ and $\operatorname{im}(B) \subseteq \ker(A)$.

    ⑧ Applying ⑤,

$$\ker(B^*) = \mathbb{R}^n \cap \ker(B^*) = [\ker(A) \oplus \operatorname{im}(A^*)] \cap \ker(B^*)$$
$$= [\ker(A) \cap \ker(B^*)] \oplus [\operatorname{im}(A^*) \cap \ker(B^*)]$$
$$= \ker(A^*A + BB^*) \oplus \operatorname{im}(A^*),$$

    where the last equality follows from ⑥ and $\operatorname{im}(A^*) \subseteq \ker(B^*)$. Alternatively, apply ⑦ with B^*, A^* in place of A, B.

    ⑨ Applying ⑤ to B^* followed by ⑧, we get

$$\mathbb{R}^n = \ker(B^*) \oplus \operatorname{im}(B) = \operatorname{im}(A^*) \oplus \ker(A^*A + BB^*) \oplus \operatorname{im}(B).$$

    ⑩ Applying ⑤ to $A^*A + BB^*$, which is self-adjoint, we see that

$$\operatorname{im}(A^*A + BB^*) = \ker(A^*A + BB^*)^\perp = \operatorname{im}(A^*) \oplus \operatorname{im}(B),$$

    where the last equality follows from ⑨. $\qquad\qquad\square$

Any two vector spaces of the same dimension are isomorphic. So saying that two vector spaces are isomorphic isn't saying very much — just that they have the same dimension. The two spaces in (2.12) are special because they are *naturally isomorphic*, i.e., if you construct an isomorphism, and the guy in the office next door constructs an isomorphism, both of you would end up with the same isomorphism, namely, the one below.

THEOREM A.3. *Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$ with $AB = 0$. Then the following spaces are naturally isomorphic*

$$\ker(A)/\operatorname{im}(B) \cong \ker(A) \cap \ker(B^*) \cong \ker(B^*)/\operatorname{im}(A^*).$$

PROOF. Let $\pi : \mathbb{R}^n \to \operatorname{im}(B)^\perp$ be the orthogonal projection of $\mathbb{R}^n$ onto the orthogonal complement of $\operatorname{im}(B)$. So any $x \in \mathbb{R}^n$ has a unique decomposition into two mutually orthogonal components

$$\begin{aligned} \mathbb{R}^n &= \operatorname{im}(B)^\perp \oplus \operatorname{im}(B), \\ x &= \pi(x) + (1 - \pi)(x). \end{aligned}$$

Let π_A be π restricted to the subspace $\ker(A)$. So any $x \in \ker(A)$ has a unique decomposition into two mutually orthogonal components

$$\begin{aligned} \ker(A) &= \big(\ker(A) \cap \operatorname{im}(B)^\perp\big) \oplus \operatorname{im}(B), \\ x &= \pi_A(x) + (1 - \pi_A)(x), \end{aligned}$$

bearing in mind that $\ker(A) \cap \operatorname{im}(B) = \operatorname{im}(B)$ since $\operatorname{im}(B) \subseteq \ker(A)$.

As π is surjective, so is π_A. Hence $\operatorname{im}(\pi_A) = \ker(A) \cap \operatorname{im}(B)^\perp$. Also, for any $x \in \ker(A)$, $\pi_A(x) = 0$ iff the component of x in $\operatorname{im}(B)^\perp$ is zero, i.e., $x \in \operatorname{im}(B)$. Hence $\ker(\pi_A) = \operatorname{im}(B)$. The first isomorphsim theorem,

$$\ker(A)/\ker(\pi_A) \cong \operatorname{im}(\pi_A) = \ker(A) \cap \operatorname{im}(B)^\perp$$

yields the required result since $\operatorname{im}(B)^\perp = \ker(B^*)$ by ③. The other isomorphism may be obtained as usual by using B^*, A^* in place of A, B. $\qquad\square$

In mathematics, *linear algebra* usually refers to a collection of facts that follow from the defining axioms of a field and of a vector space. In this regard, every single statement in Theorems A.1, A.2, A.3 is false as a statement in linear algebra — they depend specifically on our working over a subfield of $\mathbb{C}$ and are not true over arbitrary fields. For example, consider the finite field of two elements $\mathbb{F}_2 = \{0, 1\}$ and take

$$A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}.$$

Then $A^* = A = B = B^*$, and $AB = B^*A^* = A^*A = BB^* = A^*A + BB^* = 0$, which serves as a counterexample to ①, ②, ⑤, ⑦, ⑧, ⑨, ⑩, and Theorem A.3.

Appendix B. Div, Grad, Curl, and All That

We provide routine verifications of statements claimed in Sections 3 and 4.

LEMMA B.1. *Equip $L^2(V)$ and $L^2_\wedge(E)$ with the inner products in (3.1), we have*

$$\operatorname{grad}^* X(i) = -\sum_{j=1}^n \frac{w_{ij}}{w_i} X(i, j) = -\operatorname{div} X(i).$$

PROOF. The required expression follows from

$$\langle \operatorname{grad}^* X, f \rangle_V = \langle X, \operatorname{grad} f \rangle_E$$

$$= \sum\nolimits_{i<j} w_{ij} X(i,j) \operatorname{grad} f(i,j)$$

$$= \sum\nolimits_{i<j} w_{ij} X(i,j)[f(j) - f(i)]$$

$$= \sum\nolimits_{i<j} w_{ij} X(i,j) f(j) + \sum\nolimits_{i<j} w_{ij} X(j,i) f(i)$$

$$\overset{①}{=} \sum\nolimits_{j<i} w_{ji} X(j,i) f(i) + \sum\nolimits_{i<j} w_{ij} X(j,i) f(i)$$

$$\overset{②}{=} \sum\nolimits_{j<i} w_{ij} X(j,i) f(i) + \sum\nolimits_{i<j} w_{ij} X(j,i) f(i)$$

$$= \sum\nolimits_{i \neq j} w_{ij} X(j,i) f(i)$$

$$= \sum\nolimits_{i=1}^{n} w_i \Big[\sum\nolimits_{j:j \neq i} \frac{w_{ij}}{w_i} X(j,i) \Big] f(i)$$

$$\overset{③}{=} \sum\nolimits_{i=1}^{n} w_i \Big[\underbrace{\sum\nolimits_{j=1}^{n} \frac{w_{ij}}{w_i} X(j,i)}_{\operatorname{grad}^* X(i)} \Big] f(i).$$

① follows from swapping labels i and j in the first summand.
② follows from $w_{ij} = w_{ji}$.
③ follows from $X(i,i) = 0$.   □

LEMMA B.2. *Equip $L^2_\wedge(E)$ and $L^2_\wedge(T)$ with the inner products in* (3.1), *we have*

$$\operatorname{curl}^* \Phi(i,j) = \sum_{k=1}^{n} \frac{w_{ijk}}{w_{ij}} \Phi(i,j,k).$$

PROOF. The required expression follows from

$$\langle \operatorname{curl}^* \Phi, X \rangle_E = \langle \Phi, \operatorname{curl} X \rangle_T = \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) \operatorname{curl} X(i,j,k)$$

$$= \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k)[X(i,j) + X(j,k) + X(k,i)]$$

$$= \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(i,j) + \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(j,k)$$
$$+ \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(k,i)$$

$$\overset{①}{=} \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(i,j) + \sum\nolimits_{i<j<k} w_{ijk} \Phi(j,k,i) X(j,k)$$
$$+ \sum\nolimits_{i<j<k} w_{ijk} \Phi(k,i,j) X(k,i)$$

$$\overset{②}{=} \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(i,j) + \sum\nolimits_{k<i<j} w_{kij} \Phi(i,j,k) X(i,j)$$
$$+ \sum\nolimits_{i<k<j} w_{ikj} \Phi(j,i,k) X(j,i)$$

$$\overset{③}{=} \sum\nolimits_{i<j<k} w_{ijk} \Phi(i,j,k) X(i,j) + \sum\nolimits_{k<i<j} w_{kij} \Phi(i,j,k) X(i,j)$$
$$+ \sum\nolimits_{i<k<j} w_{ikj} \Phi(i,j,k) X(i,j)$$

$$= \sum\nolimits_{i<j} \Big[\Big(\sum\nolimits_{k=j+1}^{n} + \sum\nolimits_{k=1}^{i-1} + \sum\nolimits_{k=i+1}^{j-1} \Big) w_{ijk} \Phi(i,j,k) \Big] X(i,j)$$

$$= \sum_{i<j} w_{ij} \Big[\sum_{k:k\neq i,j} \frac{w_{ijk}}{w_{ij}} \Phi(i,j,k)\Big] X(i,j)$$

$$\overset{\text{④}}{=} \sum_{i<j} w_{ij} \Big[\underbrace{\sum_{k=1}^{n} \frac{w_{ijk}}{w_{ij}} \Phi(i,j,k)}_{\text{curl}^* \Phi(i,j)}\Big] X(i,j).$$

① follows from the alternating property of Φ.
② follows from relabeling j,k,i as i,j,k in the second summand and swapping labels j and k in the third summand.
③ follows from $\Phi(j,i,k)X(j,i) = \Phi(i,j,k)X(i,j)$ since both changed signs.
④ follows from $\Phi(i,j,i) = \Phi(i,j,j) = 0$. $\qquad\square$

LEMMA B.3. *The operator* $\Delta_0 = -\operatorname{div}\operatorname{grad}$ *gives us the usual graph Laplacian.*

PROOF. Let $f \in L^2(V)$. By definition,

$$\operatorname{grad} f(i,j) = \begin{cases} f(j) - f(i) & \text{if } \{i,j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

Define the adjacency matrix $A \in \mathbb{R}^{n\times n}$ by

$$a_{ij} = \begin{cases} 1 & \text{if } \{i,j\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

The gradient may be written as $\operatorname{grad} f(i,j) = a_{ij}(f(j) - f(i))$ and so

$$\begin{aligned}
(\Delta_0 f)(i) &= -[\operatorname{div}(\operatorname{grad} f)](i) = -[\operatorname{div} a_{ij}(f(j) - f(i))](i) \\
&= -\sum_{j=1}^{n} a_{ij}[f(j) - f(i)] = d_i f(i) - \sum_{j=1}^{n} a_{ij} f(j),
\end{aligned}$$

(B.1)

where for any vertex $i = 1, \ldots, n$, we define its degree as

$$d_i = \deg(i) = \sum_{j=1}^{n} a_{ij}.$$

If we regard a function $f \in L^2(V)$ as a vector $(f_1, \ldots, f_n) \in \mathbb{R}^n$ where $f(i) = f_i$ and set $D = \operatorname{diag}(d_1, \ldots, d_n) \in \mathbb{R}^{n\times n}$, then (B.1) becomes

$$\Delta_0 f = \begin{bmatrix} d_1 - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & d_2 - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \cdots & d_n - a_{nn} \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} = (D - A)f.$$

So Δ_0 may be regarded as $D - A$, the usual definition of a graph Laplacian. $\qquad\square$

THEOREM B.4. *We have that*

$$\operatorname{curl}\operatorname{grad} = 0, \qquad \operatorname{div}\operatorname{curl}^* = 0,$$

and more generally, for $k = 1, 2, \ldots,$

$$\delta_k \delta_{k-1} = 0, \qquad \delta_{k-1}^* \delta_k^* = 0.$$

PROOF. We only need to check $\delta_k \delta_{k-1} = 0$. The other relations follow from taking adjoint or specializing to $k = 1$. Let $f \in L^2_\wedge(K_{k-1})$. By (4.2) and (4.3),

$$(\delta_k \delta_{k-1} f)(i_0, \ldots, i_{k+1}) = \sum_{j=0}^{k+1} (-1)^j \delta_{k-1} f(i_0, \ldots, \widehat{i}_j, \ldots, i_{k+1})$$

$$\stackrel{\text{\textcircled{1}}}{=} \sum_{j=0}^{k+1} (-1)^j \Big[\sum_{\ell=0}^{j-1} (-1)^\ell f(i_0, \ldots, \widehat{i}_\ell, \ldots, \widehat{i}_j, \ldots, i_{k+1})$$

$$+ \sum_{\ell=j+1}^{k+1} (-1)^{\ell-1} f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1}) \Big]$$

$$= \sum_{j<\ell} (-1)^j (-1)^\ell f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1})$$

$$+ \sum_{j>\ell} (-1)^j (-1)^{\ell-1} f(i_0, \ldots, \widehat{i}_\ell, \ldots, \widehat{i}_j, \ldots, i_{k+1})$$

$$\stackrel{\text{\textcircled{2}}}{=} \sum_{j<\ell} (-1)^{j+\ell} f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1})$$

$$+ \sum_{\ell>j} (-1)^{j+\ell-1} f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1})$$

$$= \sum_{j<\ell} (-1)^{j+\ell} f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1})$$

$$- \sum_{j<\ell} (-1)^{j+\ell} f(i_0, \ldots, \widehat{i}_j, \ldots, \widehat{i}_\ell, \ldots, i_{k+1}) = 0.$$

The power of -1 in the third sum in ① is $\ell - 1$ because an argument preceding $\widehat{i}_\ell$ is omitted and so $\widehat{i}_\ell$ is the $(\ell - 1)$th argument (which is also omitted). ② follows from swapping labels j and ℓ in the second sum. $\qquad\square$

Appendix C. Calculations

We will work out the details of Example 4.1. While we have defined coboundary operators and Hodge Laplacians as abstract, coordinate-free linear operators, any actual applications would invariably involve 'writing them down as matrices' to facilitate calculations. Readers might perhaps find our concrete approach here instructive.

A simple recipe for writing down a matrix representing a coboundary operator and a Hodge Laplacian is as follows: Given an undirected graph, label its vertices and edges arbitrarily but differently for easy distinction (e.g. we used numbers for vertices and letters for edges) and assign arbitrary directions to the edges. From the graphs in Figure 6, we get the labeled directed graphs G_1 (left) and G_2 (right) in Figure 9.
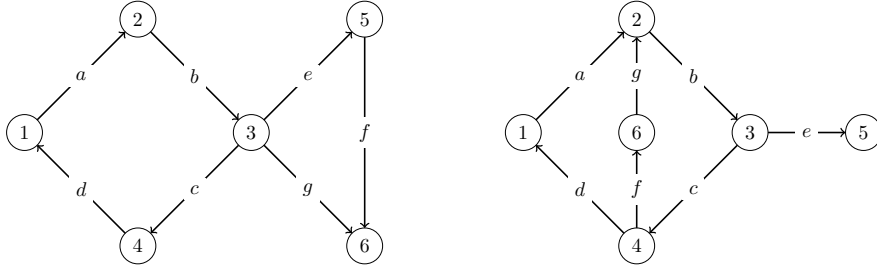

FIGURE 9. The graphs in Figure 6, with vertices and edges arbitrarily labeled and directions on edges arbitrarily assigned.

The next step is to write down a matrix whose columns are indexed by the vertices and the rows are indexed by the edges and whose (i, j)th entry is $+1$ if jth edge points into the ith vertex, -1 if jth edge points out of the ith vertex, and 0 otherwise. This matrix represents the gradient operator $\delta_0 = \mathrm{grad}$. We get

$$
A_1 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array}
\begin{array}{c} 1 \quad\; 2 \quad\; 3 \quad\; 4 \quad\; 5 \quad\; 6 \\
\left[\begin{array}{cccccc}
-1 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & -1 & 1 \\
0 & 0 & -1 & 0 & 0 & 1
\end{array} \right] \end{array}, \;
A_2 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array}
\begin{array}{c} 1 \quad\; 2 \quad\; 3 \quad\; 4 \quad\; 5 \quad\; 6 \\
\left[\begin{array}{cccccc}
-1 & 1 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & -1 & 1 & 0 & 0 \\
1 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & 0 & -1 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & -1
\end{array} \right] \end{array}
$$

for G_1 and G_2 respectively. Note that every row must contain exactly one $+1$ and one -1 since every edge is defined by a pair of vertices. This matrix is also known as a vertex-edge incidence matrix of the graph. Our choice of ± 1 for in/out-pointing edges is also arbitrary — the opposite choice works just as well as long as we are consistent throughout.

The graph Laplacians may either be computed from our definition as

$$
L_1 = A_1^* A_1 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}
\begin{array}{c} 1 \quad\; 2 \quad\; 3 \quad\; 4 \quad\; 5 \quad\; 6 \\
\left[\begin{array}{cccccc}
2 & -1 & 0 & -1 & 0 & 0 \\
-1 & 2 & -1 & 0 & 0 & 0 \\
0 & -1 & 4 & -1 & -1 & -1 \\
-1 & 0 & -1 & 2 & 0 & 0 \\
0 & 0 & -1 & 0 & 2 & -1 \\
0 & 0 & -1 & 0 & -1 & 2
\end{array} \right] \end{array},
$$

$$
L_2 = A_2^* A_2 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array}
\begin{array}{c} 1 \quad\; 2 \quad\; 3 \quad\; 4 \quad\; 5 \quad\; 6 \\
\left[\begin{array}{cccccc}
2 & -1 & 0 & -1 & 0 & 0 \\
-1 & 3 & -1 & 0 & 0 & -1 \\
0 & -1 & 3 & -1 & -1 & 0 \\
-1 & 0 & -1 & 3 & 0 & -1 \\
0 & 0 & -1 & 0 & 1 & 0 \\
0 & -1 & 0 & -1 & 0 & 2
\end{array} \right] \end{array},
$$

or written down directly using the usual definition [**21, 63**],

$$
\ell_{ij} = \begin{cases}
\deg(v_i) & \text{if } i = j, \\
-1 & \text{if } v_i \text{ is adjacent to } v_j, \\
0 & \text{otherwise.}
\end{cases}
$$

We obtain the same Laplacian matrix irrespective of the choice of directions on edges and the choice of ± 1 for in/out-pointing edges. For us there is no avoiding the gradient operators since we need them for the graph Helmholtzian below.

We may now find the eigenvalues of L_1 and L_2 and see that they are indeed the values we claimed in Example 4.1:

$$
\lambda(L_1) = (0, \; 3 - \sqrt{5}, \; 2, \; 3, \; 3, \; 3 + \sqrt{5}) = \lambda(L_2).
$$

To write down the graph Helmholtzians, we first observe that G_1 has exactly one triangle (i.e., 2-clique) whereas G_2 has none[11]. We will need to label and pick an arbitrary orientation for the triangle in G_1: We denote it as T and orient it clockwise $3 \to 5 \to 6 \to 3$. A matrix representing the operator $\delta_1 = \mathrm{curl}$ may be similarly written down by indexing the columns with edges and the rows with triangles. Here we make the arbitrary choice that if the jth edge points in the same direction as the orientation of the ith triangle, then the (i,j)th entry is $+1$ and if it points in the opposite direction, then the entry is -1. For G_1 we get

$$B_1 = T \begin{array}{c} \begin{matrix} a & b & c & d & d & e & f \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 & -1 \end{bmatrix} \end{array}.$$

Since G_2 contains no triangles, $B_2 = 0$ by definition.

We compute the graph Helmholtzians from definition,

$$H_1 = A_1 A_1^* + B_1^* B_1 = \begin{array}{c} \begin{matrix} a & b & c & d & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{matrix} \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 1 & 0 & 1 \\ -1 & 0 & -1 & 2 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 3 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 3 \end{bmatrix} \end{array}$$

$$H_2 = A_2 A_2^* + B_2^* B_2 = \begin{array}{c} \begin{matrix} a & b & c & d & d & e & f \end{matrix} \\ \begin{matrix} a \\ b \\ c \\ d \\ e \\ f \\ g \end{matrix} \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 1 \\ -1 & 2 & -1 & 0 & -1 & 0 & -1 \\ 0 & -1 & 2 & -1 & 1 & -1 & 0 \\ -1 & 0 & -1 & 2 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 2 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 2 & -1 \\ 1 & -1 & 0 & 0 & 0 & -1 & 2 \end{bmatrix} \end{array}$$

and verify that they have different spectra, as we had claimed in Example 4.1,

$$\lambda(H_1) = (0,\, 3-\sqrt{5},\, 2,\, 3,\, 3,\, 3,\, 3+\sqrt{5}) \neq (0,\, 0,\, 3-\sqrt{5},\, 2,\, 3,\, 3,\, 3+\sqrt{5}) = \lambda(H_2).$$

We now repeat the routine and convert the undirected graphs in Figure 7 into labeled directed graphs G_3 (left) and G_4 (right) in Figure 10. We label both
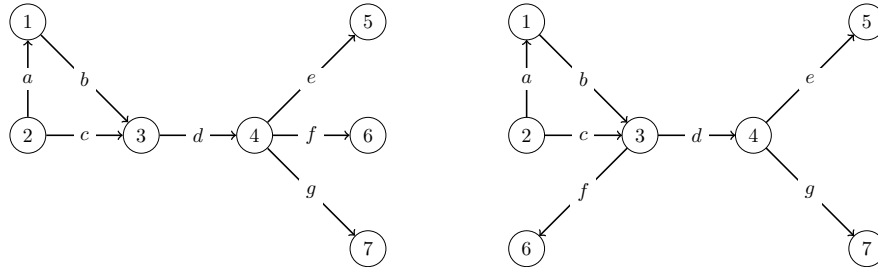

FIGURE 10. Labeled directed versions of the graphs in Figure 7.

---

[11]Those who see two triangles should note that these are really squares, or C_4's to be accurate. See also Example 3.1.

triangles in G_3 and G_4 as T and orient it clockwise $2 \to 1 \to 3 \to 2$, giving us a matrix that represents both curl operators on G_3 and G_4,

$$B_3 = B_4 = T \begin{array}{ccccccc} a & b & c & d & e & f & g \\ \left[\begin{array}{ccccccc} 1 & 1 & -1 & 0 & 0 & 0 & 0 \end{array}\right]. \end{array}$$

With these choices, we obtain the following matrix representations of the gradients, Laplacians, and Helmholtzians on G_3 and G_4,

$$A_3 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left[\begin{array}{ccccccc} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right], \end{array}$$

$$A_4 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left[\begin{array}{ccccccc} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right], \end{array}$$

$$L_3 = A_3^* A_3 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left[\begin{array}{ccccccc} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 3 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 4 & -1 & -1 & -1 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right], \end{array}$$

$$L_4 = A_4^* A_4 = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \end{array} \begin{array}{ccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ \left[\begin{array}{ccccccc} 2 & -1 & -1 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ -1 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 3 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 1 \end{array}\right], \end{array}$$

$$H_3 = A_3 A_3^* + B_3^* B_3 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{ccccccc} a & b & c & d & e & f & g \\ \left[\begin{array}{ccccccc} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 3 & -1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 2 & -1 & -1 & -1 \\ 0 & 0 & 0 & -1 & 2 & 1 & 1 \\ 0 & 0 & 0 & -1 & 1 & 2 & 1 \\ 0 & 0 & 0 & -1 & 1 & 1 & 2 \end{array}\right], \end{array}$$

$$H_4 = A_4 A_4^* + B_4^* B_4 = \begin{array}{c} \\ a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \begin{array}{ccccccc} a & b & c & d & e & f & g \\ \left[\begin{array}{ccccccc} 3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 3 & -1 & 0 & -1 & 0 \\ 0 & -1 & -1 & 2 & -1 & 1 & -1 \\ 0 & 0 & 0 & -1 & 2 & 0 & 1 \\ 0 & -1 & -1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 0 & -1 & 1 & 0 & 2 \end{array}\right] \end{array}.$$

As we intend to show that G_3 and G_4 have isospectral Hodge k-Laplacians for all k, we will also need to examine the Hodge 2-Laplacian Δ_2. Since G_3 and G_4 have no cliques of order higher than two, $\delta_k = 0$ for all $k > 2$ and in particular $\Delta_2 = \delta_1 \delta_1^*$. So the 1×1 matrix representing Δ_2 is just

$$P_3 = B_3 B_3^* = [3] = B_4 B_4^* =: P_4$$

for both G_3 and G_4.

Finally, we verify that the spectra of the Hodge k-Laplacians of G_3 and G_4 are identical for $k = 0, 1, 2$, as we had claimed in Example 4.1:

$$\lambda(L_3) = (0,\ 0.40,\ 1,\ 1,\ 3,\ 3.34,\ 5.26) = \lambda(L_4),$$
$$\lambda(H_3) = (0.40,\ 1,\ 1,\ 3,\ 3,\ 3.34,\ 5.26) = \lambda(H_4),$$
$$\lambda(P_3) = 3 = \lambda(P_4).$$

Observe that three eigenvalues of L_3, L_4, H_3, H_4 have been rounded to two decimal places — these eigenvalues have closed form expressions (zeros of a cubic polynomial) that are a hassle to typeset. So to verify that they are indeed isospectral, we check their characteristic polynomials instead, since these have integer coefficients and can be expressed exactly:

$$\det(L_3 - xI) = -21x + 112x^2 - 209x^3 + 178x^4 - 73x^5 + 14x^6 - x^7$$
$$= -x(x-3)(x-1)^2(x^3 - 9x^2 + 21x - 7) = \det(L_4 - xI),$$
$$\det(H_3 - xI) = 63 - 357x + 739x^2 - 743x^3 + 397x^4 - 115x^5 + 17x^6 - x^7$$
$$= -(x-3)^2(x-1)^2(x^3 - 9x^2 + 21x - 7) = \det(H_4 - xI).$$

COMPUTATIONAL AND APPLIED MATHEMATICS INITIATIVE, DEPARTMENT OF STATISTICS, UNIVERSITY OF CHICAGO, CHICAGO, IL 60637-1514.

E-mail address: lekheng@galton.uchicago.edu