LAPLACIANS OF CELLULAR SHEAVES
THEORY AND APPLICATIONS

JAKOB HANSEN

A DISSERTATION
in
Applied Mathematics and Computational Science

Presented to the Faculties of the University of Pennsylvania
in Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy

2020

Supervisor of Dissertation

_____

Robert W. Ghrist, Andrea Mitchell University Professor
Professor of Mathematics
Professor of Electrical and Systems Engineering

Graduate Group Chairperson

_____

Pedro Ponte Castañeda, Raymond S. Markowitz Faculty Fellow
Professor of Mechanical Engineering and Applied Mechanics

Dissertation Committee:

Robert Ghrist, Professor of Mathematics and Electrical and Systems Engineering
Philip Gressman, Professor of Mathematics
Victor Preciado, Associate Professor of Electrical and Systems Engineering

ACKNOWLEDGMENTS

This dissertation was largely written in an intense period of effort in early 2020, just before the coronavirus pandemic spread across the United States and upended everything. Now, from a vantage point in social isolation, I find myself looking back at the nearly decade-long history culminating in the production of this thesis. The influence, guidance, and leadership of many people have made this work possible.

First among these is my advisor, Robert Ghrist. His unchanging confidence in my work has been at times greater than my own, and his support and encouragement have been crucial. He has steered me in fruitful research directions, reassured me that I have created something useful, shielded me from academic politics and policies, and served as the foremost booster of my work.

The other members of my dissertation committee deserve recognition as well. It was in Victor Preciado's course on distributed systems that I learned about the graph Laplacian and discovered that it generalized to cellular sheaves. His excitement about the idea and willingness to explore applications spurred my own research into potential uses of sheaf Laplacians. Philip Gressman offered useful advice and provided moral support when needed.

The applied topology research group at Penn has been a fruitful source of ideas and advice. As fellow graduate students, Greg Henselman, Iris Yoon, Sam Carp, Huy Mai, Ann Sizemore, Darrick Lee, Hans Reiss, and Zoe Cooperband have been comrades in learning, sources of inspiration, and sounding boards for my nascent ideas. In particular, my officemate Darrick has been an ever-willing audience to my inchoate ideas, unsolved problems, and wild hopes. As postdocs in Rob's group Brendan Fong and Rachel Levanger were instrumental in cultivating my mathematical taste and exposing me to new ideas.

Others who have served as inspiration and cheerleaders for my work include Chad Giusti, Vidit Nanda, and Justin Curry. In particular, this thesis owes a huge debt to Justin's dissertation [Cur14], which is a comprehensive survey laying out a program for cellular sheaves in applied topology. Without such an excellent reference as a foundation, none of this thesis would have ever been written.

I am grateful to have learned topology at Penn from Herman Gluck, Jonathan Block, and Mona Merling, as well as geometry from Wolfgang Ziller and Davi Maximo and algebra from David Harbater. For my exposure to engineering tools and applications I thank Shuo Han and George Pappas. Moving backwards in time, I owe debts of gratitude to those who encouraged and mentored me in mathematical research before graduate school, including Rosemary Renaut, Douglas Cochran, Eric Kostelich, and Jared Culbertson. In particular, I must thank Matthias Kawski,

whose course in point set topology confirmed a decision to shift my focus from economics to mathematics.

I must also thank the contributors to Math StackExchange, MathOverflow, Wikipedia, and the nLab for the swaths of mathematical knowledge they have made easy to discover, access, and understand. Interacting with these resources has deepened my mathematical intuition and clarified countless points of confusion.

Finally, the nonmathematical support of my parents, family, friends, and community has been crucial in maintaining my sanity during the arduous process of obtaining a doctoral degree. Thanks to you all. Now, on to the math!

ABSTRACT

LAPLACIANS OF CELLULAR SHEAVES

THEORY AND APPLICATIONS

Jakob Hansen

Robert Ghrist

Cellular sheaves are a discrete model for the theory of sheaves on cell complexes. They carry a canonical cochain complex computing their cohomology. This thesis develops the theory of the Hodge Laplacians of this complex, as well as avenues for their application to concrete engineering and data analysis problems. The *sheaf Laplacians* so developed are a vast generalization of the graph Laplacians studied in spectral graph theory. As such, they admit generalizations of many results from spectral graph theory and the spectral theory of discrete Hodge Laplacians. A theory of approximation of cellular sheaves is developed, and algorithms for producing spectrally good approximations are given, as well as a generalization of the notion of expander graphs. Sheaf Laplacians allow development of various dynamical systems associated with sheaves, and their behavior is studied. Finally, applications to opinion dynamics, extracting network structure from data, linear control systems, and distributed optimization are outlined.

CONTENTS

LIST OF FIGURES

PREFACE

This dissertation is perhaps a bit out of the ordinary. It is an effort in applied mathematics, but takes a less common approach. Rather than taking a problem to be solved and focusing on a solution to that particular problem, this work begins with a general phenomenon and builds up a theory around it, working to connect this theory with real-world instantiations of the phenomenon. Rather than addressing a specific scientific question, we seek to build a better understanding of a broad class of behaviors and structures. As we build connections between theory and the world, questions from reality suggest further avenues for theoretical investigation.

Properly done, this approach to applied mathematics lifts our view to new questions rather than simply providing better, faster, or sharper answers to questions we already knew how to ask. This does not mean a disconnection from the real world, or an abdication of the responsibility to produce real results. Nor is this meant to denigrate the valuable work done in more conventional approaches to applied mathematics. Indeed, the dividing line between these two approaches is far from precise.

One might find a parallel here with Alexander Grothendieck's attitude towards problem solving in algebraic geometry and other fields of pure mathematics. Likening a problem to be solved with a nut to be opened, he described two plans of attack. One is to take a hammer and chisel and strike precisely at various points on the shell until it cracks apart. The other, his preferred method, is more expansive. As he described in *Récoltes et Semailles* (translated in [McL03]):

> I can illustrate the second approach with the same image of a nut to be opened. The first analogy that came to my mind is of immersing the nut in some softening liquid, and why not simply water? From time to time you rub so the liquid penetrates better, and otherwise you let time pass. The shell becomes more flexible through weeks and months—when the time is ripe, hand pressure is enough, the shell opens like a perfectly ripened avocado!
>
> A different image came to me a few weeks ago. The unknown thing to be known appeared to me as some stretch of earth or hard marl, resisting penetration... the sea advances insensibly in silence, nothing seems to happen, nothing moves, the water is so far off you hardly hear it... yet it finally surrounds the resistant substance.

Both approaches are important and valuable, but our approach here will fall squarely in the second category. The phenomenon we wish to understand is, broadly, the interaction between local and global information, and the mathematical

tool we will use to model it is the *sheaf*. This structure was introduced by Jean Leray during World War II to study partial differential equations, and quickly proved its utility in geometry and topology. A sheaf over a topological space describes rules for attaching data to that space in a way that allows local data to be patched together into a global picture. The prototypical example of a sheaf on a space X associates to each open set $U \subseteq X$ the set of continuous functions $U \to \mathbb{R}$. A pair of functions on U and V which agree on $U \cap V$ glue together to a unique function on $U \cup V$. This property is the essence of the local-to-global properties controlled by a sheaf.

The classical theory of sheaves has a simplified, discrete incarnation in the form of *cellular sheaves*. Rather than assigning data to open sets of a topological space, a cellular sheaf assigns data to cells of a regular cell complex, with maps specifying consistency relations. This perspective on sheaf theory was introduced by Allen Shepard [She85] and revived by Justin Curry [Cur14]. Since that time, sheaves have attracted increasing interest in the field of applied algebraic topology, in large part as a tool for developing the theory of constructions such as Mapper and persistent homology [Cur15; SMP16].

Here we will focus on a more direct sort of application of the theory of sheaves (also anticipated by Curry), as a representation of interconnected objects and systems. Cellular sheaves can represent data structures associated with networks of various kinds—for instance, social networks, networks of sensors, or more abstract networks like functional brain networks. The sheaf structure encodes information about how data should vary over a network or more general space, and homological properties of the sheaf can encode properties of that network as a system.

The theory of cellular sheaves is only one side of the story. Our goal is to combine the powerful algebraic tools of sheaves and cohomology with the tools of network science and spectral graph theory, in particular the graph Laplacian. The Laplacian operator on a graph both encodes information about the graph and provides a tool for building networked systems. Combining this toolkit with the expressive relationships represented by cellular sheaves augments the power of both sheaves and spectral graph theory. Spectral sheaf theory is no longer limited by rigid algebraic equality and can provide tools and insights even in the presence of uncertainty and noise. Conversely, the class of phenomena that can be modeled using graph Laplacians grows dramatically when we extend it to work with sheaves on graphs.

This thesis is divided into two parts: one focused on building the theory of sheaf Laplacians and another outlining areas for its application. As befits the "rising sea" analogy, the first section is longer. Once the theory is properly constructed, applications follow simply and naturally. Part I begins by defining cellular sheaves and constructing their Laplacians. This is followed by a quick review of previous

constructions on graphs that are reminiscent of cellular sheaves. We then move to a deeper investigation of the properties of sheaf Laplacians, including their spectra. This investigation specializes into a notion of approximation for cellular sheaves, which inspires a definition of expander sheaf generalizing the concept of an expander graph. The final chapter of Part I explores dynamics that can be associated with a cellular sheaf. Part II describes how many of the results obtained in the previous section can be used in practice. Sheaf cohomology, Laplacian spectra, approximation, and dynamics combine to solve problems in network science and engineering distributed systems.

Earlier versions of some of the material included in this thesis were formerly published in

> Jakob Hansen and Robert Ghrist. Toward a Spectral Theory of Cellular Sheaves. *Journal of Applied and Computational Topology* 3:4 (2019).

Some of the applications in Part II were previously described in

> Jakob Hansen and Robert Ghrist. Learning Sheaf Laplacians from Smooth Signals. IEEE International Conference on Acoustics, Sound, and Signal Processing, 2019.
>
> Jakob Hansen and Robert Ghrist. Distributed Optimization with Sheaf Homological Constraints. Allerton Conference on Communication, Control, and Computing, 2019.

Part I

THEORY

1

CELLULAR SHEAVES AND THEIR LAPLACIANS

This chapter introduces the tools we will use in our investigation of the interaction between local and global: sheaves and cosheaves over regular cell complexes. We then move to the construction of their Laplacians via discrete Hodge theory and describe a few basic properties. These definitions and constructions will carry us through the rest of our investigation. Getting them right clarifies our thinking and makes future results straightforward.

1.1 CELL COMPLEXES

Cell complexes are a familiar construction from algebraic topology; they are a class of topological spaces that can be built by attaching discs ("cells") of progressively higher dimension to a space in a controlled way. For combinatorial and computational purposes, it is useful to restrict further to *regular* cell complexes.

Definition 1.1.1. A *regular cell complex* X is a topological space partitioned into *cells* $\{X_\alpha\}_{\alpha \in P_X}$ which satisfies the following:

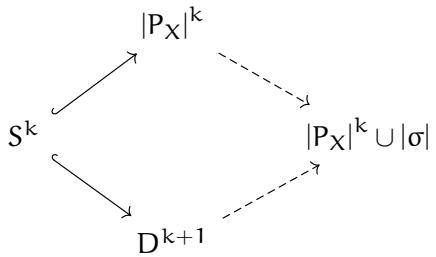
1. For each $x \in X$, every sufficiently small neighborhood of x intersects finitely many X_α.
2. For all $\alpha, \beta \in P_X$, $\overline{X_\alpha} \cap X_\beta$ is nonempty only if $X_\beta \subseteq \overline{X_\alpha}$.
3. For every X_α, there is a homeomorphism $\varphi_\alpha : D^{d_\alpha} \to \overline{X_\alpha}$ from the closed d_α-dimensional ball to the closure of X_α, which maps the interior of D^{d_α} homeomorphically onto X_α.

The second condition ensures that P_X is a partially ordered set, where $\beta \leqslant \alpha$ when $X_\beta \subseteq \overline{X_\alpha}$. This is the *face incidence poset* of X. We typically write this relation as $\beta \trianglelefteq \alpha$, and say "$\beta$ is a face of α" (or alternately, "α is a coface of β"). When there are no elements separating β and α in the face incidence poset (and $\beta \neq \alpha$), we can be more specific and write $\beta \trianglelefteq_1 \alpha$, read "$\beta$ is a codimension-one face of α." The face incidence poset is graded by the dimensions of the cells X_α. We write $\dim \alpha$ for the dimension d_α of the ball homeomorphic to X_α. The requirement that the φ_α must be homeomorphisms in condition 3 is the condition that makes X regular, and it ensures that we can reconstruct X completely from its face poset P_X, something that is not true for general cell complexes.

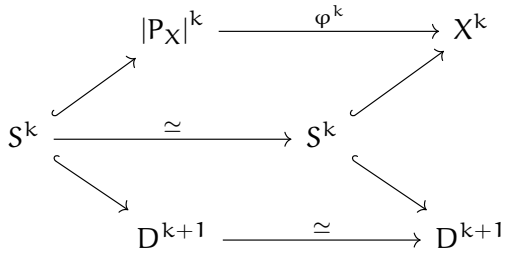
Proposition 1.1.1. *Let X be a regular cell complex, with P_X its face incidence poset. From P_X we can construct a topological space $|P_X|$ which is homeomorphic to X.*

Proof. We construct $|P_X|$ inductively by dimension, together with a homeomorphism $\varphi : |P_X| \to X$. The 0-skeleton $|P_X|^0$ consists of one point for each $\alpha \in P_X$ of dimension 0; the homeomorphism φ^0 sends $|P_X|_\alpha$ to X_α.

Once we have constructed the k-skeleton $|P_X|^k$, we can add a $(k+1)$-cell σ by looking at its faces. By the regularity condition, the faces of σ must form a homeomorphic image of S^k in $|P_X|^k$, so we have the pushout diagram



We inductively assume that we have a homeomorphism $\varphi^k : |P_X|^k \to X^k$. The same pushout diagram exists for X^k and X_σ, so we join the two diagrams:



Since this diagram commutes, there is an induced homeomorphism between $|P_X|^k \cup |\sigma|$ and $X^k \cup X_\sigma$. Repeat this process for all cells of dimension $k+1$ to extend φ^k to a homeomorphism $\varphi^{k+1} : |P_X|^{k+1} \to X^{k+1}$. □

The class of face incidence posets of regular cell complexes has a combinatorial characterization, given in [Bjö84]. To explain this, we need a few definitions.

Definition 1.1.2. A *shelling* of a simplicial complex X is an ordering $\sigma_1, \sigma_2, \ldots$ of the maximal simplices of X such that for every $k > 1$, all maximal simplices of the complex

$$X_k = \left(\bigcup_{i=1}^{k-1} \sigma_i \right) \cap \sigma_k$$

have dimension $\dim \sigma_k - 1$. That is, if we build X by gluing together the simplices σ_i in order, each new simplex glues along a boundary consisting of cells of codimension 1 in the new simplex.

Definition 1.1.3. Let P be a partially ordered set, with $\alpha \leqslant \beta$ in P. The *interval* $[\alpha, \beta] \subseteq P$ is the set of all $\gamma \in P$ with $\alpha \leqslant \gamma \leqslant \beta$.

Definition 1.1.4. Let P be a partially ordered set. The *nerve* $N(P)$ of P is a simplicial complex whose k-simplices consist of strictly increasing chains of length $k+1$ in P. Thus, the 0-simplices of $N(P)$ are the elements of P, the 1-simplices of $N(P)$ are pairs $x_1 < x_0 \in P$, and so on. Since a chain in P is simply a subset satisfying a particular condition, and a subset of a chain is a chain, this collection of subsets of P is an abstract simplicial complex.

We say that a poset is shellable if its nerve is shellable. Then the characterization of face incidence posets of regular cell complexes is as follows:

Proposition 1.1.2 ([Bjö84], Prop. 2.2)**.** *P_X is the face incidence poset of a regular cell complex if and only if, after adjoining a minimal element 0, the following hold:*

1. *Every interval $[\alpha, \beta]$ of length two has four elements.*
2. *Every interval $[0, \alpha]$ is finite and shellable.*

Condition 1 is geometric in origin. For instance, it ensures that every 1-cell has two distinct endpoints, and that there are precisely two ways to get from a 0-cell to a 2-cell by "going up" to cofaces. Condition 2 ensures that cells glue together nicely on a more global scale.

We also want a combinatorially nice class of morphisms for our complexes. For our purposes, we will call a continuous map $f : X \to Y$ between regular cell complexes a *regular cellular map* if the image of each cell X_α of X is a cell $Y_{f(\alpha)}$ of Y with $\dim Y_{f(\alpha)} \leqslant \dim X_\alpha$, and the map restricted to X_α is equivalent via some homeomorphism to the Euclidean projection $\mathbb{R}^{\dim X_\alpha} \to \mathbb{R}^{\dim Y_{f(\alpha)}}$. (This condition rules out pathological maps like space-filling curves.) Such a map induces a poset morphism $P_X \to P_Y$, which we will also call f, except in the following proposition, where they will be distinguished for clarity.

Proposition 1.1.3. *A poset morphism $\varphi : P_X \to P_Y$ of face incidence posets arises from a regular cellular map $f : X \to Y$ if and only if for all $\alpha \in P_x$*

1. $\dim \varphi(\alpha) \leqslant \dim \alpha$ *and*
2. $\varphi(\overline{\alpha}) = \overline{\varphi(\alpha)}$,

where $\overline{\alpha}$ is the downset *consisting of all elements preceding α in the partial ordering.*

Proof. First, note that $X_{\overline{\alpha}} = \overline{X_\alpha}$, for if $\beta \leqslant \alpha$, $X_\beta \subseteq \overline{X_\alpha}$. Thus, taking downsets in P_X corresponds to taking closures of cells in X.

The requirement that the image of X_α be a cell of Y ensures that a regular cellular map commutes with closures of cells. Continuity of a map f ensures that $f(\overline{X_\alpha}) \subseteq \overline{f(X_\alpha)}$. On the other hand, $\overline{X_\alpha}$ is homeomorphic to D^n and hence is compact, so its image is also compact, and hence $f(\overline{X_\alpha})$ is a closed set containing $f(X_\alpha)$, so it contains $\overline{f(X_\alpha)}$. Thus, a regular cellular map f induces a poset map satisfying the conditions (1) and (2).

Conversely, if $\varphi : P_X \to P_Y$ is a poset morphism between face incidence posets satisfying (1) and (2), we construct a map $f : X \to Y$ by induction on skeleta. The dimension condition (1) ensures that we can define f on 0-skeleta. Suppose we have a map $f^k : X^k \to Y$ which represents our poset morphism on X^k. We want to extend f^k over a $(k+1)$-cell $\sigma \in P_X$, so we consider the restriction of f^k to $\partial\sigma \subseteq X^k$. Since $\varphi(\overline{\sigma}) = \overline{\varphi(\sigma)}$, we just need to extend the map $X_{\partial\sigma} \to Y_{f(\sigma)}$ to X_σ. Since $Y_{f(\sigma)}$ is homeomorphic to a closed disk, it is contractible, so such an extension exists. Repeating this construction for all $(k+1)$-cells σ yields $f^{k+1} : X^k \to Y$, and by induction we get a map $f : X \to Y$. □

Note that the map constructed in this proof is unique up to homotopy, as there is only one homotopy class of extension of f over each cell σ.

Two classes of regular cellular maps will be important in what follows. First are *locally injective* maps. This is simple to interpret topologically—there is a neighborhood of every point on which the map is injective. To understand local injectivity combinatorially, we need to understand the way the topology of X is encoded in the face poset P_X. The subsets of X recognized by the cell structure are precisely the unions of cells, so the smallest neighborhood of a point x is the *star* of the cell containing x. The star of a cell σ, denoted $\mathrm{st}(\sigma)$, consists of all cells of which σ is a face; equivalently, this is the minimal upwards-closed set containing σ in P_X. Thus, for a cellular map $f : X \to Y$ to be locally injective, it must be the case that for each cell σ of X, $f|_{\mathrm{st}(\sigma)}$ is an injective, dimension-preserving map on the cell poset.

We will also consider *covering maps*. By similar considerations, a map $\pi : E \to X$ is a covering map if for each $\sigma \in X$, $\pi^{-1}(\mathrm{st}(\sigma))$ is a disjoint union of isomorphic copies of $\mathrm{st}(\sigma)$, each of which maps isomorphically onto $\mathrm{st}(\sigma)$.

Simplicial complexes and cubical complexes are special cases of regular cell complexes, and simplicial and cubical maps are special cases of regular cellular maps. In particular, 1-dimensional regular cell complexes are equivalent to undirected multigraphs without self-loops. We will therefore sometimes use letters such as v (for vertex) to refer to 0-cells, and e (for edge) to refer to 1-cells. In certain situations, it may be convenient to loosen the regularity requirements slightly to allow 1-cells to have attaching maps which are not homeomorphisms. This amounts

to allowing "self-loops" at vertices. While this violates the conditions above, it is usually possible (at least for 1-dimensional complexes) to extend constructions for regular cell complexes to these complexes with loops. At other times, we may wish to tighten the restrictions for 1-dimensional complexes, so that there can be at most one edge between any two vertices. This coincides with a more standard definition of graph. For the most part, however, these distinctions will not be important.

Lifting simplicial and cubical complexes to regular cell complexes allows us to perform certain constructions more readily. For instance, the product of two simplicial complexes is naturally represented by a cell complex, but requires many more simplices to realize as a simplicial complex. If X and Y are (finite) regular cell complexes, $X \times Y$ has a natural cell structure, with cells $\sigma_X \times \sigma_Y$ for σ_X a cell of X and σ_Y a cell of Y. The dimension of $\sigma_X \times \sigma_Y$ is $\dim(\sigma_X) + \dim(\sigma_Y)$, and $\sigma_X \times \sigma_Y \trianglelefteq \tau_X \times \tau_Y$ if $\sigma_X \trianglelefteq \tau_X$ and $\sigma_Y \trianglelefteq \tau_Y$. The projection maps $\pi_X : X \times Y \to X$ and $\pi_Y : X \times Y \to Y$ are regular cellular maps with respect to this cell structure. These maps send a cell $\sigma_X \times \sigma_Y$ is sent to σ_X and σ_Y, respectively.

For the remainder of this thesis, we will identify a regular cell complex X with its face incidence poset P_X, writing such things as $\sigma \in X$ to say that σ is a cell of X. The preceding arguments should be enough to convince the skeptic that no great harm is done by this abuse of notation. Working in the category of cell complexes efficiently combinatorializes all of the topology of the spaces we deal with, giving us a concrete, readily computable substrate on which to build a theory of sheaves and their Laplacians.

Additional information about regular cell complexes may be found in [LW69] and [Cur14].

1.2 CELLULAR SHEAVES

Regular cell complexes form a natural class of combinatorial spaces. These combinatorial properties make it easy to construct a tractable theory of sheaves. The shortest way to define cellular sheaves is as follows:

Definition 1.2.1. A *cellular sheaf* valued in a category $\mathcal{C}$ on a regular cell complex X is a (covariant) functor $\mathcal{F} : P_X \to \mathcal{C}$.

Let's unravel this definition a bit. We think of $\mathcal{C}$ as a data category, where objects represent spaces of data that could potentially be assigned to different cells of X. Natural candidates are such categories as the category **Set** of sets and functions, the category $\mathbf{Vect}_{\mathbf{k}}$ of $\mathbf{k}$-vector spaces and linear maps, or the category **Met** of metric spaces and continuous maps. The data of a functor $\mathcal{F}$ from P_X to $\mathcal{C}$ provides two things. First, for each cell σ of X, we get an object $\mathcal{F}(\sigma)$ of $\mathcal{C}$. Further, for every

incident pair of cells $\sigma \trianglelefteq \tau$, we get a morphism $\mathcal{F}_{\sigma \trianglelefteq \tau} : \mathcal{F}(\sigma) \to \mathcal{F}(\tau)$ in $\mathcal{C}$. The functoriality condition ensures that if $\sigma \trianglelefteq \tau \trianglelefteq \chi$, we have $\mathcal{F}_{\tau \trianglelefteq \chi} \circ \mathcal{F}_{\sigma \trianglelefteq \tau} = \mathcal{F}_{\sigma \trianglelefteq \chi}$.

We call the object $\mathcal{F}(\sigma)$ the *stalk* of $\mathcal{F}$ over σ, and the morphism $\mathcal{F}_{\sigma \trianglelefteq \tau}$ the *restriction map* from σ to τ.

One thinks of a cellular sheaf as describing data parameterized by a space, together with local consistency conditions for that data. Consistent assignments of data to stalks are known as *sections*.

Definition 1.2.2. The *sections* of a cellular sheaf $\mathcal{F}$ over an open union of cells $U \subseteq X$ are given by the limit of the diagram $(\mathcal{F}(\sigma))_{\sigma \in U}$. The *global sections* of $\mathcal{F}$ are the sections over X. The sections of $\mathcal{F}$ over U are denoted $\mathcal{F}(U)$ or $\Gamma(U; \mathcal{F})$.

In general, we will be working with data categories which are **Set**-like, where objects are sets with some extra structure and morphisms are set functions satisfying extra conditions. (In particular, these categories are complete, so that the limit defining the sections of a sheaf exists.) In this setting, a section of $\mathcal{F}$ over U is determined by giving its value at each stalk in U. Concretely, we can represent a section over U by giving some x_σ for each $\sigma \in U$, such that whenever $\sigma \trianglelefteq \tau$ with both cells in U, we have $\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma = x_\tau$. Commutativity of the diagram of stalks ensures that a section is in fact determined by its values on the stalks over cells of minimal dimension in U. Thus, for instance, a global section is determined by its values on stalks over 0-cells.

Compare the definition of a cellular sheaf with the standard definition of a presheaf on a topological space.

Definition 1.2.3. A *presheaf* valued in $\mathcal{C}$ on a topological space X is a *contravariant* functor $\mathcal{F} : O(X) \to \mathcal{C}$, where $O(X)$ is the poset of open sets in X ordered by inclusion.

Cellular sheaves seem backwards compared with presheaves on topological spaces. This is because the face incidence poset is dual to a poset of open sets. The precise statement of this correspondence relies on the *Alexandrov topology* on P_X. This is the topology generated by upwards-closed sets. Each cell σ has a minimal neighborhood, namely the star $\mathrm{st}(\sigma)$, consisting of all cells of which σ is a face.[1] We can think of a cellular sheaf on X as a presheaf defined on the cover of X given by all stars of simplices, since if $\sigma \trianglelefteq \tau$, $\mathrm{st}(\sigma) \supseteq \mathrm{st}(\tau)$. This presheaf extends to a sheaf defined on all open unions of cells.

The upshot of all of this is that we may think of a cellular sheaf on X as a genuine sheaf in two ways. One is as a sheaf on P_X with the Alexandrov topology, and the other is as a *constructible sheaf* on X, determined by its values on stars. These perspectives are developed in detail in Justin Curry's thesis [Cur14].

1 This fact also justifies the terminology "stalk," since the diagram for the colimit defining $\mathcal{F}_\sigma$ has a terminal element $\mathcal{F}(\sigma)$.

1.2.1 *Cosheaves*

Dualizing the definition of cellular sheaves gives us cellular cosheaves.

Definition 1.2.4. A *cellular cosheaf* valued in $\mathcal{C}$ on a regular cell complex X is a contravariant functor $\hat{\mathcal{F}} : P_X \to \mathcal{C}$.

Thus, the data of a cosheaf consists of *costalks* $\hat{\mathcal{F}}(\sigma)$ over cells of X along with *extension maps* $\hat{\mathcal{F}}_{\sigma \trianglelefteq \tau} : \hat{\mathcal{F}}(\tau) \to \hat{\mathcal{F}}(\sigma)$ such that $\hat{\mathcal{F}}_{\sigma \trianglelefteq \tau} \circ \hat{\mathcal{F}}_{\tau \trianglelefteq \chi} = \hat{\mathcal{F}}_{\sigma \trianglelefteq \chi}$ for $\sigma \trianglelefteq \tau \trianglelefteq \chi$.

Sheaves naturally have sections, while cosheaves naturally have *cosections*.

Definition 1.2.5. The *cosections* of a cellular cosheaf $\hat{\mathcal{F}}$ over an open union of cells $U \subseteq X$ are given by the colimit of the diagram $(\hat{\mathcal{F}}(\sigma))_{\sigma \in U}$. The *global cosections* of $\hat{\mathcal{F}}$ are the cosections over X. We will denote cosections of $\hat{\mathcal{F}}$ over U by $\hat{\mathcal{F}}(U)$ or $\hat{\Gamma}(U; \hat{\mathcal{F}})$.

Because they are constructed by colimits rather than limits, cosections (at least of cosheaves valued in **Set**-like categories) are somewhat more mysterious than sections. While sections can be thought of as lifts of an underlying complex, or as generalized sorts of functions, cosections are more like equivalence classes of data located at different places in a space. This is a somewhat nebulous interpretation which is nevertheless useful when thinking about phenomena like parallel transport or flows (see Section 1.10).

For most of this thesis, we will focus on cellular sheaves and cosheaves valued in **Vect**, although sheaves valued in **Set** will also play an important role.

Definition 1.2.6. A morphism $\varphi : \mathcal{F} \to \mathcal{G}$ between two cellular sheaves (or cosheaves) over X is a natural transformation between their functors $P_X \to \mathcal{C}$.

When we disassemble this definition, we see that a sheaf morphism $\varphi : \mathcal{F} \to \mathcal{G}$ consists of morphisms $\varphi_\sigma : \mathcal{F}(\sigma) \to \mathcal{G}(\sigma)$ subject to the condition that $\varphi_\tau \circ \mathcal{F}_{\sigma \trianglelefteq \tau} = \mathcal{G}_{\sigma \trianglelefteq \tau} \circ \varphi_\sigma$. Sheaf morphisms give a consistent way of translating local data valued in one sheaf to another sheaf. Morphisms of cosheaves are analogous; they consist of component maps $\varphi_\sigma : \hat{\mathcal{F}}(\sigma) \to \hat{\mathcal{G}}(\sigma)$ that commute with the extension maps.

Taking the global sections of a sheaf is a functorial operation. That is, we have a functor Γ from the category $\mathbf{Shv}_{\mathcal{C}}(X)$ of $\mathcal{C}$-valued sheaves to $\mathcal{C}$. A sheaf morphism $\varphi : \mathcal{F} \to \mathcal{G}$ induces a morphism $\Gamma\varphi : \Gamma(X; \mathcal{F}) \to \Gamma(X; \mathcal{G})$, and these induced morphisms behave properly with respect to composition. Global cosections of a cosheaf are similarly functorial.

Morphisms of sheaves of vector spaces have kernels and cokernels, and the category of cellular sheaves of vector spaces over X is an abelian category. Kernels and cokernels may be computed stalkwise, and it is a nice exercise in diagram chasing to show that they are in fact cellular sheaves, i. e., they are again functors $P_X \to \mathbf{Vect}$. A morphism of cellular sheaves is an epimorphism if it is stalkwise

surjective, and conversely a monomorphism if it is stalkwise injective. Naturally, a sheaf isomorphism is a sheaf morphism which is an isomorphism on all stalks.

Definition 1.2.7 (Constant sheaf)**.** Let $\mathbb{V}$ be a vector space. The *constant sheaf* on X with stalk $\mathbb{V}$ is the cellular sheaf $\underline{\mathbb{V}}$ with $\underline{\mathbb{V}}(\sigma) = \mathbb{V}$ for all σ and $\underline{\mathbb{V}}_{\sigma \trianglelefteq \tau} = \mathrm{id}_{\mathbb{V}}$ for all $\sigma \trianglelefteq \tau$.

When we work with sheaves of $\mathbf{k}$-vector spaces, the most important constant sheaf is $\underline{\mathbf{k}}$. Sections of $\mathcal{F}$ correspond with sheaf morphisms $\varphi : \underline{\mathbf{k}} \to \mathcal{F}$. To see this, observe that $\varphi_\sigma(1)$ selects an element of $\mathcal{F}(\sigma)$ for every σ, and by the commutativity condition for sheaf morphisms, we have $\mathcal{F}_{\sigma \trianglelefteq \tau} \circ \varphi_\sigma(1) = \varphi_\tau \circ \underline{\mathbf{k}}_{\sigma \trianglelefteq \tau}(1) = \varphi_\tau(1)$. Thus the elements $(\varphi_\sigma(1))_{\sigma \in X}$ determine a section of $\mathcal{F}$.[2]

1.3 SHEAF OPERATIONS

Part of the power of sheaf theory is the availability of operations that produce new sheaves from old ones. This perspective was championed by Alexander Grothendieck, who identified six important functors: two pushforward and pullback functors corresponding to every map of spaces, and the internal Hom and tensor product. The standard sheaf operations are simpler to define for cellular sheaves than for their continuous counterparts, and this simplicity will allow us to understand their interactions with spectral constructions more readily.

Definition 1.3.1 (Direct sum)**.** The *direct sum* of two sheaves $\mathcal{F}$ and $\mathcal{G}$ on a cell complex X is the sheaf $\mathcal{F} \oplus \mathcal{G}$ with stalks $(\mathcal{F} \oplus \mathcal{G})(\sigma) = \mathcal{F}(\sigma) \oplus \mathcal{G}(\sigma)$ and restriction maps $(\mathcal{F} \oplus \mathcal{G})_{\sigma \trianglelefteq \tau} = \mathcal{F}_{\sigma \trianglelefteq \tau} \oplus \mathcal{G}_{\sigma \trianglelefteq \tau}$.

Definition 1.3.2 (Tensor product)**.** The *tensor product* of two sheaves $\mathcal{F}$ and $\mathcal{G}$ on a cell complex X is the sheaf $\mathcal{F} \otimes \mathcal{G}$ with stalks $(\mathcal{F} \otimes \mathcal{G})(\sigma) = \mathcal{F}(\sigma) \otimes \mathcal{G}(\sigma)$ and restriction maps $(\mathcal{F} \otimes \mathcal{G})_{\sigma \trianglelefteq \tau} = \mathcal{F}_{\sigma \trianglelefteq \tau} \otimes \mathcal{G}_{\sigma \trianglelefteq \tau}$.

The constant sheaf $\underline{0}$ is an identity for the direct sum: for any sheaf $\mathcal{F}$, $\underline{0} \oplus \mathcal{F} \simeq \mathcal{F} \simeq \mathcal{F} \oplus \underline{0}$. Similarly, $\underline{\mathbf{k}}$ is an identity for the tensor product: $\underline{\mathbf{k}} \otimes \mathcal{F} \simeq \mathcal{F} \simeq \mathcal{F} \otimes \underline{\mathbf{k}}$.

Definition 1.3.3 (Sheaf Hom)**.** The *sheaf* Hom between two sheaves $\mathcal{F}$ and $\mathcal{G}$ on a cell complex X is the sheaf $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})$ where $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})(\sigma)$ is the space of sheaf morphisms $\mathcal{F}|_{\mathrm{st}(\sigma)} \to \mathcal{G}|_{\mathrm{st}(\sigma)}$. Since every sheaf morphism $\mathcal{F}|_{\mathrm{st}(\sigma)} \to \mathcal{G}|_{\mathrm{st}(\sigma)}$ contains the data of a sheaf morphism $\mathcal{F}|_{\mathrm{st}(\tau)} \to \mathcal{F}|_{\mathrm{st}(\tau)}$ for $\sigma \trianglelefteq \tau$, there is an obvious restriction map $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})_{\sigma \trianglelefteq \tau}$.

2 This argument also holds for sheaves valued in **Set**, with the constant sheaf with one-element stalks taking the place of $\underline{\mathbf{k}}$.

Note that if the restriction maps of $\mathcal{F}$ are invertible, a sheaf morphism on $\mathrm{st}(\sigma)$ may be given by a morphism $\mathcal{F}(\sigma) \to \mathcal{G}(\sigma)$, since we then have $\varphi_\tau = \mathcal{G}_{\sigma \trianglelefteq \tau} \circ \varphi_\sigma \circ \mathcal{F}^{-1}_{\sigma \trianglelefteq \tau}$. In this case, we can represent $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})(\sigma)$ by the space $\mathrm{Hom}(\mathcal{F}(\sigma), \mathcal{G}(\sigma))$, and the restriction maps by $\underline{\mathrm{Hom}}(\mathcal{F}, \mathcal{G})_{\sigma \trianglelefteq \tau}(\varphi) = \mathcal{G}_{\sigma \trianglelefteq \tau} \circ \varphi \circ \mathcal{F}^{-1}_{\sigma \trianglelefteq \tau}$.

Definition 1.3.4 (Pullback)**.** The *pullback* of a sheaf $\mathcal{F}$ on a cell complex Y over a cellular map $f : X \to Y$ is the sheaf $f^*\mathcal{F}$ on X with stalks $f^*\mathcal{F}(\sigma) = \mathcal{F}(f(\sigma))$. The restriction maps are given by $f^*\mathcal{F}_{\sigma \trianglelefteq \tau} = \mathcal{F}_{f(\sigma) \trianglelefteq f(\tau)}$.

Definition 1.3.5 (Pushforward)**.** The *pushforward* of a sheaf $\mathcal{F}$ on a cell complex X over a cellular map $f : X \to Y$ is the sheaf $f_*\mathcal{F}$ on Y with stalks $f_*\mathcal{F}(\sigma) = \lim_{\sigma \trianglelefteq f(\sigma')} \mathcal{F}(\sigma')$. The restriction maps are induced by the fact that if $\sigma \trianglelefteq \tau$, then the diagram for $f_*\mathcal{F}(\sigma)$ contains the diagram for $f_*\mathcal{F}(\tau)$ and hence $f_*\mathcal{F}(\sigma)$ is the apex of a cone over the diagram for $f_*\mathcal{F}(\tau)$ and we get an induced map $f_*\mathcal{F}_{\sigma \trianglelefteq \tau} : f_*\mathcal{F}(\sigma) \to f_*\mathcal{F}(\tau)$.

When f is a locally injective cellular map, the relevant diagram for computing $f_*\mathcal{F}(\sigma)$ consists of several components, each with a single initial object, and so $f_*\mathcal{F}(\sigma)$ is canonically isomorphic to the direct sum of these initial objects. This characterization will be useful for defining weighted pushforwards.

The other two standard sheaf operations, the pushforward with compact supports $f_!$ and the derived pullback with compact supports $f^!$, also have descriptions in terms of cellular sheaves, but they will be less useful in building a spectral theory, so we will elide their definitions.

1.4 SHEAF COHOMOLOGY

A regular cell complex X is naturally partitioned into cells of different dimensions. This allows us to construct cellular homology and cohomology theories. Cellular sheaf cohomology is a form of cellular cohomology with locally varying coefficient spaces. Given a cellular sheaf of vector spaces $\mathcal{F}$ on X, we can produce a complex of $\mathcal{F}$-valued cochains on X, graded by cell dimension.

Definition 1.4.1 (Cochains)**.** The space of $\mathcal{F}$-valued k-cochains on X is the vector space

$$C^k(X; \mathcal{F}) = \bigoplus_{\dim \sigma = k} \mathcal{F}(\sigma).$$

With an appropriately defined coboundary, these cochains form a complex. To construct the coboundary, we will need one additional gadget, a *signed incidence relation* on X. This is a function $[\bullet : \bullet] : P_X \times P_X \to \{0, \pm 1\}$ which satisfies

1. $[\sigma : \tau] \neq 0$ if and only if $\sigma \leqslant_1 \tau$.
2. For any 1-cell χ with faces σ, τ, $[\sigma : \chi][\tau : \chi] = -1$.
3. For any $\sigma \leqslant \tau$, $\sum_{\chi \in P_\chi} [\sigma : \chi][\chi : \tau] = 0$.

The number $[\sigma : \tau]$ represents whether the assigned orientations of σ and τ agree.

Given a signed incidence relation on X, we define the coboundary $\delta^k : C^k(X; \mathcal{F}) \to C^{k+1}(X; \mathcal{F})$ by

$$\delta^k|_{\mathcal{F}(\sigma)} = \sum_{\dim \tau = k+1} [\sigma : \tau] \mathcal{F}_{\sigma \leqslant \tau}$$

or equivalently

$$(\delta^k x)_\tau = \sum_{\dim \sigma = k} [\sigma : \tau] \mathcal{F}_{\sigma \leqslant \tau}(x_\sigma).$$

It is straightforward to check that $\delta^k \circ \delta^{k-1} = 0$. We evaluate at a $(k-1)$-cochain supported on a single cell σ:

$$\begin{aligned}(\delta^k \delta^{k-1} x_\sigma)_\tau &= \sum_{\dim \chi = k} [\chi : \tau] \mathcal{F}_{\chi \leqslant \tau}((\delta^{k-1} x_\sigma)_\chi) \\ &= \sum_{\dim \chi = k} [\chi : \tau] \mathcal{F}_{\chi \leqslant \tau}([\sigma : \chi] \mathcal{F}_{\sigma \leqslant \chi} x_\sigma) \\ &= \sum_{\dim \chi = k} [\chi : \tau][\sigma : \chi] \mathcal{F}_{\sigma \leqslant \tau}(x_\sigma) = 0.\end{aligned}$$

Thus we have a complex

$$C^0(X; \mathcal{F}) \xrightarrow{\delta^0} C^1(X; \mathcal{F}) \xrightarrow{\delta^1} C^2(X; \mathcal{F}) \xrightarrow{\delta^2} \cdots$$

of vector spaces associated with $\mathcal{F}$. The cohomology of this complex is the *sheaf cohomology* of X with coefficients in $\mathcal{F}$, denoted $H^\bullet(X; \mathcal{F})$. A sheaf morphism $\varphi : \mathcal{F} \to \mathcal{G}$ induces a map of complexes $C\varphi : C^\bullet(X; \mathcal{F}) \to C^\bullet(X; \mathcal{G})$, and hence a map $H\varphi : H^\bullet(X; \mathcal{F}) \to H^\bullet(X; \mathcal{G})$.

When the sheaf under consideration is the constant sheaf with stalk $\mathbb{V}$, the corresponding cochain complex is simply the complex of $\mathbb{V}$-valued cellular cochains of X. This is the tensor product $C^\bullet(X; \mathbf{k}) \otimes \mathbb{V}$, and hence the cohomology is isomorphic to $H^\bullet(X; \mathbf{k}) \otimes \mathbb{V}$.

How should we interpret sheaf cohomology? The first place to start is in degree zero. This has a simple interpretation: $H^0(X; \mathcal{F})$ is isomorphic to the space of global sections of $\mathcal{F}$. To see this, recall first that a global section of $\mathcal{F}$ is uniquely

determined by its values on stalks over 0-cells, that is, by a 0-cochain $x \in C^0(X;\mathcal{F})$. A 0-cochain x is a section precisely when, for every 1-cell e with incident 0-cells u, v, we have $\mathcal{F}_{v \trianglelefteq e} x_v = \mathcal{F}_{u \trianglelefteq e} x_u$. On the other hand, x is in $H^0(X;\mathcal{F})$ if $\delta^0 x = 0$, so we must have $\sum_v [v : e] \mathcal{F}_{v \trianglelefteq e} x_v = 0$ for all 1-cells e. A 1-cell has two faces, say v and u, and by property (2) of the signed incidence relation, we have $[v : e] = -[u : e]$, which implies that $\delta^0 x = 0$ if and only if $\mathcal{F}_{u \trianglelefteq e} x_u = \mathcal{F}_{v \trianglelefteq e} x_v$ for every $e = u \sim v$.

One way to approach higher sheaf cohomology is from the derived functor perspective, or more simply, through the long exact sequence for sheaf cohomology. A sheaf morphism $\varphi : \mathcal{F} \to \mathcal{G}$ induces maps $C^k(X;\mathcal{F}) \to C^k(X;\mathcal{G})$ that commute with the coboundary, and hence gives a functorial homomorphism $H^k\varphi : H^k(X;\mathcal{F}) \to H^k(X;\mathcal{G})$. A short exact sequence of sheaves

$$0 \to \mathcal{E} \to \mathcal{F} \to \mathcal{G} \to 0$$

is exact on stalks, and hence induces a short exact sequence of complexes

$$0 \to C^\bullet(X;\mathcal{E}) \to C^\bullet(X;\mathcal{F}) \to C^\bullet(X;\mathcal{G}) \to 0.$$

The snake lemma then gives a long exact sequence for sheaf cohomology:

$$\cdots \to H^{k-1}(X;\mathcal{G}) \to H^k(X;\mathcal{E}) \to H^k(X;\mathcal{F}) \to H^k(X;\mathcal{G}) \to H^{k+1}(X;\mathcal{E}) \to \cdots$$

Let A be a subcomplex of X, with inclusion map i. If $\mathcal{F}$ is a sheaf on X, consider the sheaf $i_* i^* \mathcal{F}$ on X. For cells σ of X which are in A, the stalks are the same as those of $\mathcal{F}$, while for cells not in A, the stalks are the zero vector space. There is a natural sheaf morphism $\mathcal{F} \to i_* i^* \mathcal{F}$, given by the identity on stalks over cells in A, and the unique morphism $\mathcal{F}(\sigma) \to 0$ over cells not in A. Since this sheaf morphism is stalkwise surjective, it is an epimorphism, so we can extend this map to a short exact sequence

$$0 \to \ker(i_* i^*) \to \mathcal{F} \to i_* i^* \mathcal{F} \to 0.$$

The sheaf $\ker(i_* i^*)$ has stalks zero over cells in A, and stalks equal to those of $\mathcal{F}$ otherwise. We can interpret cochains of $i_* i^* \mathcal{F}$ as cochains of $\mathcal{F}$ supported on A, while cochains of $\ker(i_* i^*)$ are cochains of $\mathcal{F}$ vanishing on A. The corresponding cohomology spaces may be seen as computing relative cohomology, and we will write

$$H^\bullet(X;\ker(i_* i^*)) = H^\bullet(X, A;\mathcal{F})$$

and

$$H^\bullet(X; i_* i^* \mathcal{F}) = H^\bullet(A; \mathcal{F}),$$

so we have a long exact sequence

$$\cdots \to H^{k-1}(A; \mathcal{F}) \to H^k(X, A; \mathcal{F}) \to H^k(X; \mathcal{F}) \to H^k(A; \mathcal{F}) \to H^{k+1}(X, A; \mathcal{F}) \to \cdots$$

Returning to the case of a general short exact sequence $0 \to \mathcal{E} \to \mathcal{F} \to \mathcal{G} \to 0$, the connecting map $H^0(X; \mathcal{G}) \to H^1(X; \mathcal{E})$ gives an obstruction to lifting sections of $\mathcal{G}$ to sections of $\mathcal{F}$. In the relative cohomology case, this means that $H^1(X, A; \mathcal{F})$ is a space containing obstructions to extending sections of $\mathcal{F}$ defined on A to all of X.

All of this machinery naturally dualizes to cosheaves of vector spaces, giving spaces of cosheaf-valued chains $C_\bullet(X; \hat{\mathcal{F}})$, boundary maps $\partial_k : C_k(X; \hat{\mathcal{F}}) \to C_{k-1}(X; \hat{\mathcal{F}})$, and *cosheaf homology* $H_\bullet(X; \hat{\mathcal{F}})$. For a cosheaf $\hat{\mathcal{F}}$ on a cell complex X, $H_0(X; \hat{\mathcal{F}})$ is isomorphic to the space of cosections of $\hat{\mathcal{F}}$. There is likewise a cosheaf homology long exact sequence corresponding to a short exact sequence of cosheaves.

The standard definition of sheaf cohomology treats it as the derived functor of the global sections functor [KS90; GM03]. Fortunately, it can be shown that for cellular sheaves, the cohomology theory defined above is equivalent to this more sophisticated definition. This was perhaps first proved by Allen Shepard [She85], reiterated by Justin Curry [Cur14], and independently discovered by Everitt and Turner [ET15]. However, the construction of sheaf Laplacian operators will make essential use of this canonical model for the cohomology of a cellular sheaf, and thus is somewhat difficult to extend to other sheaf-theoretic settings common in algebraic topology and homological algebra.

1.5 WEIGHTED SHEAVES

Our data category needs one more piece of structure before we can construct Laplacian operators for cellular sheaves. We need to lift from mere vector spaces to vector spaces that carry an inner product. For a field $\mathbf{k} = \mathbb{R}$ or $\mathbb{C}$, we consider the category $\mathbf{Hilb}_\mathbf{k}$ of Hilbert spaces (i. e.complete inner product spaces) over $\mathbf{k}$.

Definition 1.5.1. A *weighted cellular sheaf* over a cell complex X is a cellular sheaf valued in $\mathbf{Hilb}_\mathbf{k}$.

The introduction of an inner product extends a sheaf from an algebraic structure centered on strict equality to something more quantitative. Given a weighted cellular sheaf, we can ask about sizes of cochains and coboundaries, and give a

precise meaning to the question of "how far" a 0-cochain is from defining a global section.

Categorically, passing to $\mathbf{Hilb}_{\mathbf{k}}$ from $\mathbf{Vect}_{\mathbf{k}}$ turns our data category into a *dagger category*. This means we have a contravariant endofunctor $\dagger : \mathbf{Hilb}_{\mathbf{k}} \to \mathbf{Hilb}_{\mathbf{k}}^{\mathrm{op}}$, which is the identity on objects and satisfies $\dagger \circ \dagger = \mathrm{Id}$.[3] This dagger endofunctor is precisely the operation of taking the adjoint of a bounded linear map between Hilbert spaces.[4] From a categorical perspective, dagger categories enable us to single out a class of *unitary* morphisms, those (iso)morphisms f such that $f^{\dagger} = f^{-1}$. This gives a finer notion of equivalence for objects of a dagger category: we care not just about isomorphism but about *unitary* isomorphism. In $\mathbf{Hilb}_{\mathbb{k}}$, unitary isomorphisms are precisely the unitary linear transformations—those which preserve the inner product.

This extends to the categorical operations used when working with sheaves. The limit used to define the space of sections of a cellular sheaf is unique up to isomorphism. However, when we add the dagger categorical structure of $\mathbf{Hilb}_{\mathbf{k}}$, we want a finer invariant. The appropriate notion is that of a *dagger limit*, recently studied by Heunen and Karvonen [HK19]. The dagger limit of a diagram is unique up to unitary isomorphism. Unfortunately, dagger limits in $\mathbf{Hilb}_{\mathbf{k}}$ are somewhat more rare than limits in $\mathbf{Vect}_{\mathbf{k}}$. In particular, $\mathbf{Hilb}_{\mathbf{k}}$ does not have most dagger pullbacks. Since the diagram for computing the global sections of a sheaf over a graph with two vertices and a single edge is a pullback diagram, as shown in Figure 1.1, this fact is somewhat disheartening. Category theory does not give us a canonical inner product structure on the space of sections of a weighted sheaf. We will have to make some choices.

Fortunately, the extra structure afforded by sheaves over cell complexes offers us a couple of reasonably canonical choices for the weighted space of global sections of a weighted cellular sheaf. One is simply to view the space of global sections of a sheaf $\mathcal{F}$ as a subspace of the space of assignments to all stalks. That is, the inner product of two sections $x, y \in \Gamma(X;\mathcal{F})$ is given by $\sum_{\sigma \in X} \langle x_\sigma, y_\sigma \rangle$. Alternately, we can view sections as a subspace of $C^0(X;\mathcal{F})$, and let $\langle x, y \rangle = \sum_{\dim \sigma = 0} \langle x_\sigma, y_\sigma \rangle$. In general, the latter choice seems to be most useful, as it corresponds with the inner product on H^0 arising from discrete Hodge theory of the cellular cochain complex.

The dagger structure affords a way to dualize a sheaf into a cosheaf with the same stalks. Simply apply the dagger functor, changing $\mathcal{F}$ into a contravariant

3 The requirement that $\dagger$ be the identity on objects makes this definition less than ideal from the categorical point of view, which prefers to look only at isomorphisms of objects rather than equality. It is best to think of a dagger category as another kind of structure like a category, rather than a piece of extra categorical data. (See the discussion in [Bär15].)

4 Unfortunately, everywhere else in this thesis, the dagger symbol $\dagger$ will represent not the dagger operation on **Hilb** but the operation of taking the Moore-Penrose pseudoinverse of a linear map between Hilbert spaces. Let the reader beware.

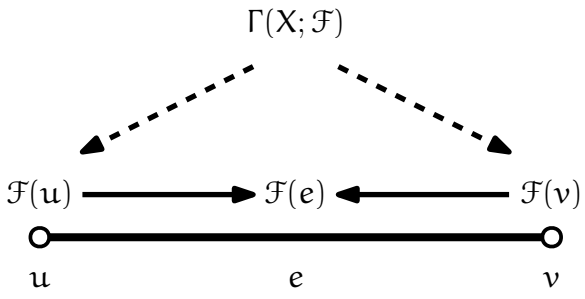


Figure 1.1: A cellular sheaf whose space of global sections is a pullback.

functor $\mathcal{F}^*$. When stalks are finite dimensional, this cosheaf is isomorphic to the cosheaf $\hat{\mathcal{F}} = \mathrm{Hom}(\bullet, \mathbf{k}) \circ \mathcal{F}$, although the particular isomorphism depends on the inner product structure.

1.6 DISCRETE HODGE THEORY

Hodge theory on a Riemannian manifold selects canonical representatives for de Rham cohomology classes by using the exterior derivative and its adjoint to construct the Hodge Laplacians $(d + d^*)^2$. A similar construction was introduced by Beno Eckmann in the 1940s to study the homology of simplicial complexes [Eck45]. This *discrete Hodge theory* is useful in approximating Hodge theory of manifolds [Dod76; Man08; Kal15], and also in discrete data analysis, with applications to problems like consistent ranking [Jia+11; Xu+17]. The construction of discrete Hodge Laplacians can be carried out for any chain complex of Hilbert spaces.

Consider a chain complex $C^\bullet$ of finite-dimensional Hilbert spaces. The coboundary map δ has an adjoint δ^*, with which we can construct the discrete Hodge Laplacian $\Delta = (\delta + \delta^*)^2$. Since δ is graded of degree $+1$ and δ^* is graded of degree -1, and since $\delta^2 = (\delta^*)^2 = 0$, Δ is graded of degree 0. Its individual graded parts are $\Delta^k = (\delta^k)^*\delta^k + \delta^{k-1}(\delta^{k-1})^*$. The key fact about the Hodge Laplacians is that their kernel computes the cohomology of $C^\bullet$.

Theorem 1.6.1. *The space* C^k *has an orthogonal decomposition* $C^k = \operatorname{im}\delta^{k-1} \oplus \ker\Delta^k \oplus \operatorname{im}(\delta^k)^*$, *and* $\ker\Delta^k \simeq H^k(C^\bullet)$.

Proof. First note that the orthogonal decomposition of C^k implies that $\ker\Delta^k$ is in fact isomorphic to the k-th cohomology of $C^\bullet$. This is because $\ker\delta^k$ is the orthogonal complement of $\operatorname{im}(\delta^k)^*$, so that $\ker\delta^k = \operatorname{im}\delta^{k-1} \oplus \ker\Delta^k$. Therefore, $\ker\Delta^k$ is the orthogonal complement of $\operatorname{im}\delta^{k-1}$ in $\ker\delta^k$. But then $\ker\Delta^k \simeq \ker\delta^k/\operatorname{im}\delta^{k-1}$.

It is straightforward to see that $\operatorname{im}\delta^{k-1} \perp \operatorname{im}(\delta^k)^*$, since for any $x \in C^{k-1}$ and $y \in C^{k+1}$, $\langle \delta^{k-1}x, (\delta^k)^* y\rangle = \langle \delta^k\delta^{k-1}x, y\rangle = 0$. As a result, if $\Delta^k x = 0$, then both $(\delta^k)^*\delta^k x = 0$ and $\delta^{k-1}(\delta^{k-1})^* x = 0$ (since the images of these two operators are orthogonal). But this implies that $\delta^k x = 0$, so $x \perp \operatorname{im}(\delta^k)^*$; similarly, $x \perp \operatorname{im}\delta^{k-1}$. Thus $\ker\Delta^k$ is orthogonal to both $\operatorname{im}(\delta^k)^*$ and $\operatorname{im}\delta^{k-1}$, and we have $C^k = \operatorname{im}\delta^{k-1} \oplus \operatorname{im}(\delta^k)^* \oplus \ker\Delta^k$. □

We call the elements of $\ker\Delta^k$ *harmonic* k*-cochains*, and denote the space of such cochains $\mathcal{H}^k(C^\bullet)$.

1.7 SHEAF LAPLACIANS

A cellular sheaf $\mathcal{F}$ with values in $\mathbf{Hilb}_\mathbf{k}$ has a complex of cochains, which can be given a canonical inner product using the orthogonal direct sum. The construction of the discrete Hodge Laplacian then applies directly. We get the *sheaf Laplacians* $\Delta^k_\mathcal{F}$, with $\ker\Delta^k_\mathcal{F} = \mathcal{H}^k(X;\mathcal{F}) \simeq H^k(X;\mathcal{F})$.

The degree-0 sheaf Laplacian is analogous to the graph Laplacian. Its kernel is a canonical representation for the space of global sections of $\mathcal{F}$. Just as the kernel of the graph Laplacian consists of locally constant functions on the vertices of a graph, the kernel of the degree-0 sheaf Laplacian consists of locally *consistent* cochains valued in the vertex stalks of a sheaf. By analogy with the graph Laplacian L_G, we will frequently denote the degree-0 sheaf Laplacian of a sheaf $\mathcal{F}$ by $L_\mathcal{F}$.

It is common in work on the spectral theory of simplicial complexes [HJ13; Par13; Ste13] to consider not only the full Hodge Laplacians, but also their separate upwards and downwards components $\Delta^k_+ = (\delta^k)^*\delta^k$ and $\Delta^k_- = \delta^{k-1}(\delta^{k-1})^*$. The up-Laplacian Δ^k_+ is also called the *coboundary* Laplacian, and its complementary down-Laplacian Δ^k_- is the *boundary* Laplacian. Most results in spectral simplicial theory deal with up-Laplacians, since these are more directly analogous to the graph Laplacian.

The direct sum structure of the spaces $C^k(X;\mathcal{F}) = \bigoplus_{\dim\sigma=k}\mathcal{F}(\sigma)$ means that we can—and usually should—think of the sheaf coboundary maps and Laplacians as block matrices. The connectivity of the underlying cell complex is encoded in the sparsity pattern of these matrices. For the degree-0 Laplacian, the blocks are fairly simple to describe. The diagonal blocks are maps $\mathcal{F}(v) \to \mathcal{F}(v)$ for 0-cells $v \in X$, and we have $L_\mathcal{F}[v,v] = \sum_{v \trianglelefteq e} \mathcal{F}^*_{v\trianglelefteq e}\mathcal{F}_{v\trianglelefteq e}$. The off-diagonal blocks are maps $\mathcal{F}(v) \to \mathcal{F}(u)$, with $u, v \trianglelefteq e$ for some 1-cell e, and the blocks are given by $L_\mathcal{F}[v,u] = -\mathcal{F}^*_{v\trianglelefteq e}\mathcal{F}_{u\trianglelefteq e}$. Compare this with the structure of the graph Laplacian, with $L[v,v] = d_v = \sum_{v\trianglelefteq e} w_e$ and $L[v,u] = -w_e$ for $u, v \trianglelefteq e$. A comparison of the graph Laplacian and the sheaf Laplacian is in Figure 1.2.

$$(Lx)_v = \sum_{u \sim v} w_{uv}(x_v - x_u)$$
$$(L_{\mathcal{F}}x)_v = \sum_{u,v \trianglelefteq e} \mathcal{F}^*_{v \trianglelefteq e}(\mathcal{F}_{v \trianglelefteq e}x_v - \mathcal{F}_{u \trianglelefteq e}x_u)$$

$$L_G = \begin{bmatrix} w_{uv} & -w_{uv} \\ -w_{uv} & w_{uv} \end{bmatrix} \qquad L_{\mathcal{F}} = \begin{bmatrix} \mathcal{F}^*_{v \trianglelefteq e}\mathcal{F}_{v \trianglelefteq e} & -\mathcal{F}^*_{v \trianglelefteq e}\mathcal{F}_{u \trianglelefteq e} \\ \mathcal{F}^*_{u \trianglelefteq e}\mathcal{F}_{v \trianglelefteq e} & \mathcal{F}^*_{u \trianglelefteq e}\mathcal{F}_{u \trianglelefteq e} \end{bmatrix}$$

$$\langle x, L_G x\rangle = \sum_{u \sim v} w_{uv}(x_v - x_u)^2$$
$$\langle x, L_{\mathcal{F}} x\rangle = \sum_{u,v \trianglelefteq e} \|\mathcal{F}_{v \trianglelefteq e}x_v - \mathcal{F}_{u \trianglelefteq e}x_u\|^2$$

Figure 1.2: Comparing the graph Laplacian and the degree-0 sheaf Laplacian

The Laplacians $\Delta^k_{\mathcal{F}}$ are positive semidefinite operators, and hence define quadratic forms $E^k(x) = \langle x, \Delta^k_{\mathcal{F}} x\rangle$ on $C^k(X;\mathcal{F})$. For $k = 0$ we have

$$E^0(x) = \langle x, L_{\mathcal{F}} x\rangle = \langle \delta^0 x, \delta^0 x\rangle = \|\delta^0 x\|^2,$$

which measures how far a 0-cochain x is from defining a section of $\mathcal{F}$.

Spectral graph theory typically concerns itself as much with *adjacency matrices* as with Laplacians. These are determined by the graph Laplacian via the formula $A = D - L$, where D is the diagonal degree matrix, equal to the diagonal of L. We can similarly define an adjacency matrix for sheaves on graphs, by letting $A_{\mathcal{F}} = \mathrm{diag}(L) - L_{\mathcal{F}}$, where $\mathrm{diag}(L)$ consists of the diagonal blocks of L. We can think of $A_{\mathcal{F}}$ as also defining a quadratic form on $C^0(X;\mathcal{F})$, with

$$\langle x, A_{\mathcal{F}} x\rangle = \sum_{u,v \trianglelefteq e} \langle \mathcal{F}_{u \trianglelefteq e}x_u, \mathcal{F}_{v \trianglelefteq e}x_v\rangle.$$

These sheaf adjacency matrices are more difficult to interpret than sheaf Laplacians. One problem is that a sheaf adjacency matrix does not uniquely determine its corresponding Laplacian. Another is that there is not a clear way to extend adjacency matrices to sheaves on higher-dimensional complexes. We will mostly ignore sheaf adjacency matrices, except in the case of matrix-weighted graphs, where Laplacians and adjacency matrices do contain equivalent information.

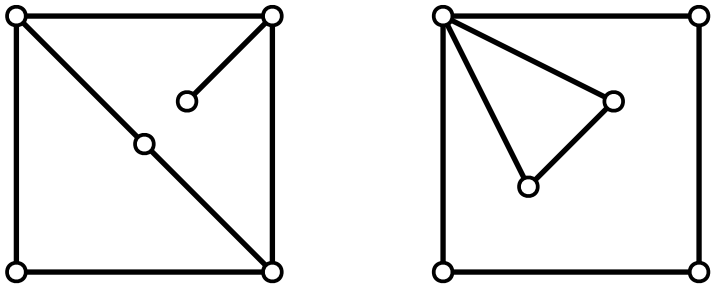


Figure 1.3: Two nonisomorphic graphs with the same Laplacian spectrum

1.7.1 *Can one hear the shape of a sheaf?*

A famous question in the spectral theory of Laplacian operators was raised by Mark Kac [Kac66]: "Can one hear the shape of a drum?" That is, from the Dirichlet spectrum of the Laplacian operator of a planar domain (i. e., the set of frequencies at which it would naturally vibrate as a drumhead), can one extract the shape of the domain? This raises the question of whether there exist distinct planar regions which are *isospectral* for the Laplacian.[5]

By analogy, we might also ask "can one hear the shape of a graph?"—that is, whether there exist nonisomorphic graphs which have the same Laplacian spectrum. There exist nonisomorphic graphs which are isospectral for the Laplacian; one example is shown in Figure 1.3. The Laplacians of both graphs have eigenvalues $\{0, 3-\sqrt{5}, 2, 3, 3, 3+\sqrt{5}\}$. More generally, we might ask if we can hear the shape of a sheaf, or to what extent the spectrum of the sheaf Laplacian determines the sheaf. This question is a superset of the question for graphs. Any graph G on n vertices can be seen as a sheaf on K_n, the complete graph on n vertices, by using some inclusion map $i : G \to K_n$ and the pushforward $i_*\underline{\mathbb{R}}$ of the constant sheaf. Two such graphs G and H are isomorphic if and only if there exist inclusions $i : G \to K_n$ and $j : H \to K_n$ such that $i_*\underline{\mathbb{R}} \simeq j_*\underline{\mathbb{R}}$. Thus, the negative answer for graphs implies a negative answer for sheaves: the spectrum of $L_{\mathcal{F}}$ does not determine $\mathcal{F}$, even up to sheaf isomorphism combined with an automorphism of the underlying graph.

The indeterminacy goes even further. Graph Laplacians are in one-to-one correspondence with (weighted) vertex-labeled graphs, so even if the spectrum does not determine the graph, the Laplacian operator itself does. However, the same is not true for cellular sheaves on graphs. Consider the two sheaves displayed in Figure 1.4.

5 For convex regions with analytic boundary, the answer is "yes, one can hear the shape of a drum" but there do exist non-convex polygonal regions with isospectral Laplacians.

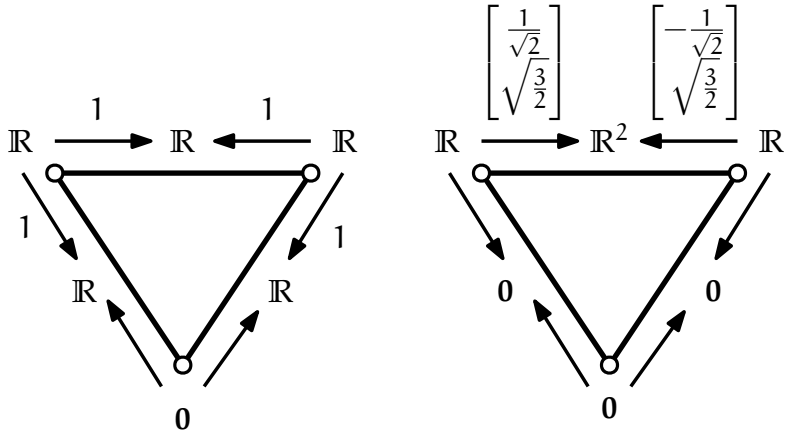


Figure 1.4: Two nonisomorphic sheaves with the same degree-0 Laplacian

Their coboundary maps have matrices

$$\delta_{\mathcal{F}} = \begin{bmatrix} 1 & -1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \delta_{\mathcal{G}} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ \sqrt{3} & -\sqrt{3} \end{bmatrix}.$$

But we have

$$L_{\mathcal{F}} = \delta_{\mathcal{F}}^* \delta_{\mathcal{F}} = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} = \delta_{\mathcal{G}}^* \delta_{\mathcal{G}} = L_{\mathcal{G}}.$$

So the answer is that not only is it not typically possible to hear the shape of a sheaf, it is not even possible to *see* the shape of a sheaf: we cannot always extract the isomorphism class of a sheaf on a graph from its (degree 0) Laplacian operator.

1.7.2 *Normalized Sheaf Laplacians*

In spectral graph theory, it is common to use a normalized version of the graph Laplacian, particularly when working with irregular graphs. The normalized Laplacian of a graph is typically defined as $\mathcal{L} = D^{-1/2} L D^{-1/2}$, where D is the diagonal matrix of vertex degrees. This normalization ensures that $\mathcal{L}$ remains symmetric and that the entries on the diagonal of $\mathcal{L}$ are all 1.

We would like to do something similar for sheaf Laplacians. Perhaps the obvious thing to do in order to normalize a degree-o sheaf Laplacian is to use the block diagonal degree matrix $D_{\mathcal{F}}$, with $D_{\mathcal{F}}[v, v] = L_{\mathcal{F}}[v, v]$ and $D_{\mathcal{F}}[u, v] = 0$ for $u \neq v$.

This matrix is symmetric and positive semidefinite, so it has a symmetric positive semidefinite pseudoinverse $D_{\mathcal{F}}^{\dagger}$, which further has a square root $D_{\mathcal{F}}^{\dagger/2}$, so we might define $\mathcal{L}_{\mathcal{F}} = D^{\dagger/2} L_{\mathcal{F}} D^{\dagger/2}$. It is not immediately clear how to extend this normalization operation to higher-degree Laplacians, nor is its meaning evident.

Instead, we turn to an approach used by Horak and Jost to define normalized Hodge Laplacians for simplicial complexes [HJ13]. The idea is that normalization is a property of the interactions of the sheaf restriction maps and the inner products on stalks, not merely a property of a matrix. An inkling of this approach appears when one notes that the matrix $D^{-1/2} L D^{-1/2}$ is similar to the matrix $D^{-1} L$, which is self-adjoint with respect to the inner product $\langle x, y \rangle = x^{\mathsf{T}} D y$. Thus, we can view the normalized graph Laplacian as reweighting each vertex proportionally to its degree.

Motivated by this perspective on the normalized graph Laplacian, we make the following definition.

Definition 1.7.1. A weighted cellular sheaf $\mathcal{F}$ defined on a regular cell complex X is *normalized* if for each cell σ of X, the restricted coboundary map $\delta|_{\mathcal{F}(\sigma)}$ is unitary on the orthogonal complement of its kernel. That is, for any $x, y \in \mathcal{F}(\sigma) \cap \ker(\delta|_{\mathcal{F}(\sigma)})^{\perp}$, we have $\langle x, y \rangle = \langle \delta x, \delta y \rangle$.

Lemma 1.7.1. *Let $\mathcal{F}$ be a weighted cellular sheaf on a finite-dimensional regular cell complex X. There exists a normalized sheaf $\tilde{\mathcal{F}}$ which is isomorphic to $\mathcal{F}$ (but not necessarily unitarily so).*

Proof. We construct $\tilde{\mathcal{F}}$ by changing only the inner products; this ensures that $\tilde{\mathcal{F}} \simeq \mathcal{F}$. Suppose X has dimension n. Then the normalization condition is trivially satisfied for all cells σ of dimension n. We recursively define the inner products on lower-dimensional stalks. Suppose σ has dimension k and $\tilde{\mathcal{F}}$ satisfies the normalization condition on all stalks of dimension greater than k. Let $\Pi_{\sigma}^{\ker}$ be the orthogonal projection $\mathcal{F}(\sigma) \to \mathcal{F}(\sigma) \cap \ker \delta$, and $\Pi_{\sigma} = \mathrm{id} - \Pi_{\sigma}^{\ker}$ its complementary projection. We then define a normalized inner product $\langle \cdot, \cdot \rangle_{\sigma}^{\sim}$ in terms of the already-normalized inner product on C^{k+1} and the current inner product on $\mathcal{F}(\sigma)$ by $\langle x, y \rangle_{\sigma}^{\sim} = \langle \delta \Pi_{\sigma} x, \delta \Pi_{\sigma} y \rangle + \langle \Pi_{\sigma}^{\ker} x, \Pi_{\sigma}^{\ker} y \rangle$. This is clearly an inner product, since the first term is nondegenerate for $x \perp \ker \delta$ and the second term is nondegenerate for $x \in \ker \delta$. Further, it obviously satisfies the normalization condition. Iterating this operation on cells of progressively lower dimension yields a normalized sheaf. □

When $\mathcal{F}$ is a sheaf on a graph, this normalization procedure requires only one step. The new inner product on a vertex stalk $\mathcal{F}(v)$ is, for $x \perp \ker \delta|_{\mathcal{F}(v)}$, given by $\langle x, y \rangle_{v}^{\sim} = \langle \delta x, \delta y \rangle = \langle x, D_v y \rangle$. The inner product on $C^0(X; \mathcal{F})$ is then $\langle x, y \rangle^{\sim} = \langle x, (D + \mathrm{id}_{\ker D}) y \rangle$. The adjoint of δ with respect to this inner product

is $(D + \mathrm{id}_{\ker D})^{-1}\delta^T = (D^\dagger + \mathrm{id}_{\ker D})\delta^T$. Thus, in terms of the original basis for $C^0(X;\mathcal{F})$, the Laplacian of the normalization is $L_{\tilde{\mathcal{F}}} = (D^\dagger + \mathrm{id}_{\ker D})\delta^T\delta = D^\dagger L_{\mathcal{F}}$. (The term $\mathrm{id}_{\ker D}$ vanishes because $\ker D \subseteq \ker L_{\mathcal{F}}$, which is orthogonal to $\operatorname{im} L_{\mathcal{F}}$.) If we change to an orthonormal basis by applying the transformation $D^{1/2} + \mathrm{id}_{\ker D}$ we get a symmetric matrix $D^{\dagger/2} L_{\mathcal{F}} D^{\dagger/2}$, the analogue of the formula for the normalized graph Laplacian.

1.8 INFINITE-DIMENSIONAL LAPLACIANS

The definitions given so far are adapted to sheaves of finite dimensional vector spaces over complexes with finitely many cells. These restrictions are not essential, but removing them adds certain complications. On a complex with infinitely many cells, there are multiple complexes of cochains:

$$\begin{aligned} C_c^\bullet(X;\mathcal{F}) &= \bigoplus_{\dim\sigma=\bullet} \mathcal{F}(\sigma) \\ C^\bullet(X;\mathcal{F}) &= \prod_{\dim\sigma=\bullet} \mathcal{F}(\sigma) \\ L^2C^\bullet(X;\mathcal{F}) &= \left\{ x \in C^\bullet(X;\mathcal{F}) : \sum_{\dim\sigma=\bullet} \|x_\sigma\| < \infty \right\}. \end{aligned}$$

For finite cell complexes, these are all equal, but in general, we only have inclusions

$$C_c^\bullet(X;\mathcal{F}) \subseteq L^2C^\bullet(X;\mathcal{F}) \subseteq C^\bullet(X;\mathcal{F}),$$

inducing morphisms on cohomology groups. Hodge theory requires an inner product, and hence is naturally defined on $L^2C^\bullet$ and (given some further conditions) computes $L^2H^\bullet(X;\mathcal{F})$. However, sections of a cellular sheaf correspond to $H^0(X;\mathcal{F})$, so that $\ker\Delta^0$ is no longer necessarily isomorphic to the space of global sections. However, the induced map $\ker\Delta^0 = L^2H^0(X;\mathcal{F}) \to H^0(X;\mathcal{F})$ is injective, since $\delta x = 0$ implies consistency across every 1-cell in X. While compactly supported sections are also in L^2H^0, sections supported everywhere on an infinite complex often will not be. Thus, the kernel of $\Delta^0_{\mathcal{F}}$ might be trivial even though $\mathcal{F}$ has nontrivial global sections.

The question of when the Hodge Laplacians of a complex of Hilbert spaces computes the L^2 cohomology of the complex was addressed in [BL92]. Some sufficient conditions include finite generation of the cohomology, which ensures that the coboundary map is a Fredholm operator. More generally, if δ and δ^* have closed images, the space of harmonic cochains will be isomorphic to the L^2 cohomology of the complex.

If we want to study the spectral properties of the Hodge Laplacians, we generally want them to be bounded linear operators. This holds if the coboundary maps are bounded.

Proposition 1.8.1. *Let $\mathcal{F}$ be a cellular sheaf of Hilbert spaces on a regular cell complex X. The following conditions are sufficient for δ^k to be a bounded linear operator:*

1. *There exists some M_k such that for every cell σ of dimension k, $\|\mathcal{F}_{\sigma \trianglelefteq \tau}\| \leqslant M_k$ for all $\sigma \trianglelefteq_1 \tau$.*
2. *Every k-dimensional cell of X has at most d^k cofaces of dimension $k+1$.*
3. *Every $(k+1)$-dimensional cell of X has at most d_{k+1} faces of dimension k.*

These conditions amount to a sort of uniform boundedness for both restriction maps of $\mathcal{F}$ and the cell complex structure of X.

Proof. Let $x \in L^2C^k(X;\mathcal{F})$. Then we have

$$\begin{aligned}
\|\delta^k x\|^2 &= \sum_{\dim \tau = k+1} \|(\delta^k x)_\tau\|^2 \\
&\leqslant \sum_{\dim \tau = k+1} \left(\sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\| \right)^2 \\
&\leqslant \sum_{\dim \tau = k+1} \left(\sum_{\sigma \trianglelefteq_1 \tau} M_k \|x_\sigma\| \right)^2 \\
&\leqslant M_k^2 \sum_{\dim \tau = k+1} d_{k+1} \sum_{\sigma \trianglelefteq_1 \tau} \|x_\sigma\|^2 \\
&= M_k^2 d_{k+1} \sum_{\dim \sigma = k} \sum_{\sigma \trianglelefteq_1 \tau} \|x_\sigma\|^2 \\
&\leqslant M_k^2 d_{k+1} d^k \sum_{\dim \sigma = k} \|x_\sigma\|^2 = M_k^2 d_{k+1} d^k \|x\|^2.
\end{aligned}$$

□

The spectral theory of bounded self-adjoint operators is relatively straightforward; there is no need to fuss about domains or essential self-adjointness. The spectrum of a bounded self-adjoint operator consists entirely of *approximate eigenvalues*: those λ for which there exists a sequence of unit vectors $\{x_i\}$ with $\|\Delta^k x_i - \lambda x_i\|$ converging to zero.

Even nicer than bounded operators are compact operators. These are operators which can be approximated by finite-rank operators. The spectrum of a compact operator consists only of eigenvalues, but these will in general accumulate at zero.

Proposition 1.8.2. *Let $\mathcal{F}$ be a sheaf of Hilbert spaces on a regular cell complex X. The following conditions are sufficient for δ^k to be a compact linear operator:*

1. *For every k-dimensional cell σ with $\sigma \trianglelefteq_1 \tau$, the restriction map $\mathcal{F}_{\sigma \trianglelefteq \tau}$ is compact.*
2. $\displaystyle\sum_{\dim \sigma = k} \sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau}\| < \infty.$

Proof. Since all restriction maps are compact, we may approximate them by sequences $\mathcal{F}^i_{\sigma \trianglelefteq \tau}$ of finite-rank operators. Combine these into an approximation $(\delta^k)^i$ of δ^k. Fix an ordering of $(k+1)$-cells of X, and let the operator $P^i : C^{k+1}(X;\mathcal{F}) \to C^{k+1}(X;\mathcal{F})$ be an orthogonal projection onto the subspace spanned by stalks over cells of index at most i. $P^i(\delta^k)^i$ is a finite-rank operator and

$$\|P^i(\delta^k)^i - \delta^k\| \leqslant \sum_{j>i} \sum_{\sigma \trianglelefteq_1 \tau_j} \|\mathcal{F}_{\sigma \trianglelefteq \tau_j}\| + \sum_{j \leqslant i} \sum_{\sigma \trianglelefteq \tau_j} \|\mathcal{F}^i_{\sigma \trianglelefteq \tau_j} - \mathcal{F}_{\sigma \trianglelefteq \tau_j}\|.$$

The right hand side goes to zero as i goes to infinity, so δ^k is approximated by a sequence of finite rank operators. □

Infinite graphs can be helpful in understanding finite graphs. For instance, every d-regular graph has the infinite d-ary tree as its universal cover. A study of the spectrum of this tree allows us to extract bounds on the spectrum of d-regular graphs. However, many of the same issues arise with operators on infinite graphs: for Laplacian and adjacency operators to be bounded, vertex degrees must be uniformly bounded, and compactness requires a summability condition on weights.

For the most part, the results in this thesis will be restricted to Laplacians of sheaves of finite-dimensional Hilbert spaces over finite cell complexes. Most results will generalize to the infinite dimensional case for compact Laplacians, although in this situation there is not in general a smallest nonzero eigenvalue, making several results trivial.

1.9 WEIGHTED CONSTANT SHEAVES

Perhaps the most obvious example of a weighted cellular sheaf is the constant sheaf $\underline{\mathbb{R}}$ on a complex X, with the standard inner product on each stalk. If we take its Laplacians, we get the unweighted graph Laplacian and the higher discrete Hodge Laplacians of the complex.

What about weighted graphs? The term "weighted sheaf" may have given it away already: weighted graphs correspond to constant sheaves $\underline{\mathbb{R}}$ with a custom inner product on each edge stalk $\mathcal{F}(e) = \mathbb{R}$. This is equivalent to specifying a positive

real number (a weight) for each edge. More generally, for a constant sheaf $\underline{\mathbb{V}}$, we should be able to choose any inner product we like on each stalk. We can always rewrite the restriction maps in terms of a standard orthonormal basis and obtain a unitarily isomorphic sheaf, but it is important to be aware that the restriction maps may no longer be represented by identity matrices. For instance, the entries $\mathcal{F}^*_{v \trianglelefteq e}\mathcal{F}_{u \trianglelefteq e}$ of the degree-0 sheaf Laplacian will not be copies of the identity map, since $\mathcal{F}^*_{v \trianglelefteq e}$ depends on the inner products on $\mathcal{F}(e)$ and $\mathcal{F}(v)$. However, assuming that we stick with the standard inner product on vertex stalks, these blocks will still be symmetric and positive definite. Indeed, if we want to think of sections of the constant sheaf $\underline{\mathbb{R}^n}$ as locally constant 0-chains in the standard basis, we should not perturb the weights on vertex stalks, and only modify it on edges. This is the typical case for the study of weighted graphs and matrix-weighted graphs.

1.10 VECTOR BUNDLES

Sheaves of vector spaces where all restriction maps are invertible are "locally constant," since they are isomorphic to a constant sheaf on each sufficiently small open set. The classical Riemann-Hilbert correspondence states that, on a sufficiently nice category of base spaces, locally constant sheaves are equivalent to vector bundles with a flat connection. (The correspondence is as follows: the kernel of the connection operator acting on the sheaf of sections of the vector bundle is a locally constant sheaf, and taking the tensor product of a locally constant sheaf with the sheaf of functions yields a sheaf which is the sheaf of sections of a vector bundle, with an obvious flat connection.) Locally constant sheaves are also known as *local systems*, and correspond with representations of the fundamental groupoid of the base space.

A cellular sheaf $\mathcal{F}$ on a cell complex X with invertible restriction maps corresponds to a genuine locally constant sheaf $|\mathcal{F}|$ on the geometric realization $|X|$. Local sections of $\mathcal{F}$ over open collections U of cells are in one-to-one correspondence with local sections of $|\mathcal{F}|$ over $|U|$. If $|X|$ is, for instance, a complex manifold, the Riemann-Hilbert correspondence gives a vector bundle $E \to X$ together with a flat connection $\nabla : \Gamma(E) \to \Gamma(E) \otimes \Omega^1(X)$ whose kernel (as a sheaf morphism) is $|\mathcal{F}|$. The cellular sheaf $\mathcal{F}$ has extra data, though, in the form of the canonical cochain complex $C^\bullet(X;\mathcal{F})$. It is useful to think of δ as a discrete approximation to the connection ∇. The stalks of the sheaf are analogous to the fibers of E, while the cell complex structure provides a sort of analogue for the de Rham complex $\Omega^\bullet(X)$. Just as a flat vector bundle has a twisted de Rham complex $\Gamma(E) \otimes \Omega^\bullet(X)$, the sheaf cochain complex is a twisted version of the cellular cochain complex. The flatness condition $\nabla^2 = 0$ corresponds to the cochain complex condition $\delta^2 = 0$. Even when the technical conditions for the Riemann-Hilbert correspondence do not obtain,

it is useful to think of a cellular sheaf with invertible restriction maps as a vector bundle with a flat connection. We will call such a sheaf a *discrete vector bundle*.

One can equally well describe a local system by a locally constant cosheaf; in the cellular case, switching between these two representations corresponds to a dualizing operation that takes the inverse of each restriction map, rather than the linear dual. For weighted sheaves and cosheaves, these two operations coincide when all restriction maps are unitary.

We will call a weighted cellular sheaf with all restriction maps in some linear group G a G-bundle. But in fact this is not quite general enough to capture some desirable objects, at least for compact groups. We might expect a "weighted $SO(1)$-bundle" on a graph to be the same as a weighted constant sheaf. But given inner products on the stalks, there is only one possibility for each map, and no variation in weights is possible. Instead, we introduce a slightly more complex definition, where restriction maps are scalar multiples of maps in G. However, these scalar multiples must satisfy some consistency relations.

For each cell σ we specify a weight or scaling factor $\alpha_\sigma > 0$, and require that for $\sigma \trianglelefteq \tau$, $\mathcal{F}_{\sigma \trianglelefteq \tau} = \frac{\alpha_\tau}{\alpha_\sigma} \rho_{\sigma \trianglelefteq \tau}$, with $\rho_{\sigma \trianglelefteq \tau} \in G$. These weights α_σ are essentially the same as the weightings for simplicial complexes described by Horak and Jost [HJ13]. We can think of this as a uniform scaling of a standard inner product on the stalks of $\mathcal{F}$. It is common, particularly when working with graphs, to weight all vertices equally, although this is not necessary, and one way to view the normalized graph Laplacian is as giving each vertex the weight $\alpha_v = \sqrt{d_v}$.

This definition is somewhat opaque in its motivation. One justification is that without a canonical basis, inner products are only meaningful in relation to maps in or out of a vector space. Scaling an inner product on stalks of a sheaf is meaningless except for its effect on the restriction maps. Another justification is more direct: we want weighted $O(n)$-bundles to correspond with connection graphs (see Chapter 2.3), which requires us to be able to assign scalar weights to edges along with orthogonal restriction maps.

The dual descriptions of local systems in terms of sheaves and cosheaves have different advantages. The sheaf formulation is naturally adapted to understanding (flat) sections of a bundle, while the cosheaf formulation more naturally represents parallel transport and holonomy. If $\mathcal{F}$ is a vector bundle on a cell complex X as represented by a sheaf, with $\hat{\mathcal{F}}$ its cosheaf representation, we think of $C^0(X;\mathcal{F})$ as representing not-necessarily-flat sections, while the isomorphic space $C_0(X;\hat{\mathcal{F}})$ consists of formal sums of elements of fibers of the bundle. $C^1(X;\mathcal{F})$ then corresponds to something like bundle-valued 1-forms, and $C_1(X;\hat{\mathcal{F}})$ corresponds to something like plans for parallel transport. A 0-cochain of $\mathcal{F}$ is flat if it is in $\ker \delta^0$, while two elements of $C_0(X;\hat{\mathcal{F}})$ are connected by parallel transport if their difference is in $\operatorname{im} \partial_0$. Moving to weighted sheaves and cosheaves, the sheaf coboundary map

gives us information about how far a bundle section is from being flat, while the cosheaf boundary tells us about how difficult it is to parallel transport one vector to another. This is related to the discussion of effective resistance and conductance in Chapter 3.2.

2

A PANOPLY OF SHEAF-LIKE CONSTRUCTIONS ON GRAPHS

The utility of annotating graphs to represent pairwise relationships between nodes has long been evident. Work done in a variety of fields has led to the development of many constructions which are similar in spirit to, or even special cases of, sheaves on graphs. The material in this chapter serves as a review of previous literature related to the spectral theory of cellular sheaves, as well as a source of inspiration for those looking to find discrete sheaves that arise in other mathematical contexts.

2.1 SIGNED GRAPHS

The term "signed graph" was introduced by Frank Harary in the 1950s [Har53; CH56] to describe a formalism for representing social networks. The construction is simple: A signing of a graph G consists of a choice of sign $\sigma_e \in \mathbb{Z}^\times = \{1, -1\}$ for each edge e of G. Because signs take values in a group, there is a natural way to give a sign to each path: simply take the product of signs of all edges in the path. Following earlier notions for triangles developed by Fritz Heider, Harary called a cycle *balanced* if its sign was positive. The graph G is then balanced if every cycle is balanced.

In sheaf theoretic terms, a signed graph is equivalent to an $O(1)$-bundle on a graph. A cycle is balanced if it has trivial holonomy, and the signed graph is balanced if its corresponding sheaf is isomorphic to the constant sheaf.

In a social network with positive edges corresponding to friendships and negative edges corresponding to enmities, we might expect this balance condition to hold. After all, "the enemy of my enemy is my friend" (and also "the friend of my friend is my friend"). Global balance occurs in a social network if there is a consistent partition of individuals into two mutually antagonistic friend groups. The expectation that social networks will tend to be balanced in this way can be used to define systems of dynamics for social networks [AKR06].

2.2 VOLTAGE AND GAIN GRAPHS

Voltage graphs and gain graphs are synonymous terms for a generalization of signed graphs to labels taking values in any group. Let G be an oriented graph, and let H be a group. A gain graph over G with values in H is an assignment of

some $g_e \in H$ to each edge e of G. The group element g_e is associated with the positive orientation of the edge e; when we use the edge in reverse, the group element is inverted. (This distinction is not necessary in signed graphs because every element of $O(1)$ is its own inverse.) The *gain* along a path is the product of the g_e for the edges traversed by the path. This product must be taken in order and respect the orientation of edges.

A sheaf-like way to think of a gain graph is as a principal H-bundle over a graph G. That is, take a set A with a free and transitive H-action ρ, and construct a sheaf of sets on G by letting each stalk be equal to A, with the restriction maps for an oriented edge $e = u \to v$ be given by $\mathcal{F}_{u \trianglelefteq e} = \rho(g_e)$, $\mathcal{F}_{v \trianglelefteq e} = \mathrm{id}$.

In the context of gain graphs, one often thinks about assignments to vertices, i. e.functions $x : V(G) \to H$. In particular, there is the special case of "satisfying" assignments, those for which $x_v = g_e x_u$ for any oriented edge $e = u \to v$. These are the sections of the associated sheaf of sets. More generally, we may ask how many edges are satisfied for a given assignment, that is, for how many edges the alignment condition $x_v = g_e x_u$ holds. This question underlies various certain discrete structures in physics, such as spin glasses and the Ising model [Zas02].

Associated with a gain graph is its set of balanced cycles. These have the property that they "glue together," that is, if a theta-shaped subgraph has two balanced cycles, the third is also balanced. Abstract properties of gain graphs have been thoroughly investigated by Thomas Zaslavsky [Zas89; Zas91; Zas18], including the definition of associated matroids generalizing the cycle matroid of a graph. One such matroid is the *frame matroid*, whose independent sets are those sets of edges whose connected components are either simply connected or consist of a single unbalanced cycle.

Gain graphs also arise in the analysis of flows on a "network with gains" [Jew62], where each edge of a network admits some input flow and outputs a scaled version of that input. The gain group is in this case either $\mathbb{R}^+$ or $\mathbb{R}^\times$, depending on if flows can be turned negative. For the obvious inclusions of these groups into $GL(1)$, we get (co)sheaves valued in **Vect** with one-dimensional stalks. The cosheaf boundary matrices turn out to be the constraint matrices for flow optimization problems over networks with gains.

Gain graphs (typically with finite gain groups) are used in topological graph theory to describe covering spaces of graphs, and in this context are called *voltage graphs*. The idea is that the group elements g_e are H-valued voltages, where we do not require Kirchoff's laws to hold. Given a voltage graph G with voltages in a discrete group H, we can construct a covering space $\pi : E \to G$ whose fiber $\pi^{-1}(v)$ over every vertex v is H, with an edge between $h_v \in \pi^{-1}(v)$ and $h_u \in \pi^{-1}(u)$ for an oriented edge $e = u \sim v$ if $g_e h_u = h_v$. This realizes the principal H-bundle defined by G as a covering space of G. More generally, if A is an H-set, we can

perform the same construction with fibers A; this is frequently done for S_n-valued voltage graphs with A an n-element set.

The application to topological graph theory is as follows: given a voltage graph G embedded in a topological surface Σ, there exists a *derived surface* Σ^α together with an embedding of the induced covering E of G in Σ^α. Some careful analysis of graphs that arise as covering spaces of very simple graphs then allows us to construct interesting embeddings of these graphs. A thorough discussion of voltage graphs in the context of topological graph theory may be found in [GT87].

2.3 CONNECTION GRAPHS

The graph Laplacian, in various normalizations, can function as a discrete approximation to the Laplace-Beltrami operator of a Riemannian manifold. This is the basis for dimensionality reduction methods like Laplacian eigenmaps [BN03] and diffusion maps [CL06], which use eigenvectors of Laplacians to give nice coordinate functions for data points sampled from a submanifold of Euclidean space.

There are other Laplacian-like operators on Riemannian manifolds, including the connection Laplacian Δ_∇, which is the degree-0 Hodge Laplacian for the complex of TM-valued differential forms. That is, if we view the Levi-Civita connection ∇ as a map $\Gamma(M;TM) \to \Omega^1(M) \otimes \Gamma(M;TM)$, the connection Laplacian is $\Delta_\nabla = \nabla^*\nabla$.[1] Singer and Wu showed that a graph Laplacian-like construction can approximate the connection Laplacian on a manifold embedded in Euclidean space [SW12]. They called the matrices so produced *graph connection Laplacians*. These matrices have a block structure with $n \times n$ blocks. Each diagonal block is a scalar multiple of the identity, while off-diagonal blocks are scalar multiples of orthogonal matrices. These graph connection Laplacians can be used to construct an analogue of the diffusion maps technique, called *vector diffusion maps*, taking into account the holonomy of parallel transport of tangent vectors when distances between points are calculated.

A graph connection Laplacian is associated to and defines a *connection graph*, where each oriented edge $u \to v$ of a graph is assigned both a positive real weight w_{uv} and an orthogonal matrix $\rho_{uv} \in O(n)$. The oppositely oriented edge has the same weight and inverse rotation $\rho_{vu} = \rho_{uv}^*$. The connection Laplacian has diagonal blocks $L[u,u] = (\sum_{u\sim v} w_{uv})\, I$ and off-diagonal blocks $L[u,v] = -w_{uv}\rho_{uv}$.

This definition should remind the reader of the data of an $O(n)$-bundle on a graph, since it is precisely that. Given a connection graph, we construct a sheaf with all stalks $\mathbb{R}^n$ whose restriction maps are either the identity or ρ_{uv}, depending on the orientation. The inner product on the edge stalk is w_{uv} times the standard

1 This definition is really that of the *Bochner Laplacian*, but this is equivalent to the connection Laplacian up to a sign. See [Nic07, ch. 10].

inner product on $\mathbb{R}^n$. Under this correspondence, the graph connection Laplacian is the same as the sheaf Laplacian, and the sheaf coboundary is a discrete analogue of the connection. The relationship with vector bundles was noted by Gao, Brodzki, and Mukherjee [GBM19], who gave an explanation of graph connection Laplacians as Hodge Laplacians associated to a twisted cochain complex. The construction in terms of the cochain complex of a cellular sheaf is more suitable, however, because it both acknowledges the freedom to choose a basis for the edge stalks and allows for an extension to higher-dimensional base spaces.

Connection graphs have attracted attention for use in problems other than manifold learning, particularly *synchronization*. This is a class of data analysis problem in which a collection of parameters must be retrieved from (noisy) knowledge of pairwise relationships between the parameters. Perhaps the canonical example is recovering orientations in $O(2)$ for a collection of camera angles given knowledge of the pairwise rotations. When this data is noisy, eigenvectors of the corresponding connection Laplacian serve as good approximate solutions [Sin11]. More generally, given this information we can try to group objects into consistent clusters, as in [CK15] and [GBM19], or try to determine which observed pairwise relationships are likely to be incorrect via a sparsification-like algorithm [ZKC14]. An interesting application of these tools arises in [HLW19], where eigenvectors of a graph connection Laplacian are used to find the correct orientations for pieces of a jigsaw puzzle.

There has also been interest in connection graphs as an avenue for "high-dimensional" generalization of spectral graph theory. Rather than raising the dimension of the base space as in spectral simplicial theory, one raises the dimensionality of the data associated with the graph. This program has produced a Cheeger inequality [BSS13] and results on random connection graphs [EKW15].

2.4 HORIZONTAL DIFFUSION MAPS

Another extension of diffusion maps methods is the *horizontal diffusion maps* algorithm recently introduced by Tingran Gao [Gao19]. This algorithm is motivated by a synchronization-like problem in shape matching. Consider a collection of shapes F_i, together with maps defining a matching between points of pairs F_i, F_j. The canonical example is a collection of bones from related species of animal, where we have computed the "best" homeomorphism between the surfaces of the bones for pairs of closely related animals. We wish to understand the global structure of this collection of shapes, in order to do things like consistent segmentation or classification.

We think of the shapes F_i as fibers of a bundle over a manifold representing the space of species. The matchings between shapes determine a connection on

the fiber bundle. The horizontal diffusion maps algorithm constructs a discrete approximation to a "horizontal" Laplacian operator defined on such a bundle. Given a Riemannian fiber bundle $\pi : E \to X$, we sample points from X to give vertices x_i of a graph G, and sample a set of points $\{y_{ik}\} = Y_i$ in the fiber E_{x_i} over each x_i. We compute a matching between probability distributions on the sets Y_i and Y_j by allowing these distributions to diffuse under parallel transport on X. Taking the union of all the Y_i, these matchings give weighted edges between points in Y_i and points in Y_j, combining into a weighted graph H with a map $H \to G$.

We think of H as a sort of discrete fiber bundle over G, and taking the pushforward of the (weighted) constant sheaf on Y gives us a sheaf representation of this fiber bundle. The Laplacian of the pushforward sheaf is what Gao terms the *graph horizontal Laplacian*, on which the horizontal diffusion maps algorithm acts by normalizing and taking eigenvectors. In the context of the "bundle of bones" discussed above, these can be used to produce low-dimensional embeddings or clusterings of the set of specimens that respect the consistency of the surface matchings, or to produce consistent partitions or landmarks on the bone surfaces.

2.5 MATRIX-WEIGHTED GRAPHS

Another way of extending spectral graph theory to higher-dimensional data is by assigning matrix-valued weights to the edges of a graph. That is, rather than assigning a positive real number to each edge, we assign an $n \times n$ positive semidefinite matrix to each edge. The corresponding Laplacian and adjacency matrices are easy to construct: if W_{uv} is the weight matrix for the edge $u \sim v$, the adjacency matrix A is a block matrix with $n \times n$ block entries $A[u, v] = W_{uv}$. The Laplacian matrix has the same block structure, with diagonal blocks $L[u, u] = \sum_{u \sim v} W_{uv}$ and off-diagonal blocks $L[u, v] = -W_{uv}$.

These constructions have been considered under the name "matrix-weighted graph," with connections to problems in engineering. For instance, consensus on matrix-weighted graphs can produce nontrivial formations of autonomous agents in the plane [Tri+18], or help understand the synchronization behaviors of certain electrical circuits [Tun16]. Others have also happened upon this natural construction, and have defined a notion of effective resistance for such graphs [ABRK19].

A positive definite matrix defines an inner product on $\mathbb{R}^n$, so if all weights are strictly positive, we can view a matrix-weighted graph as defining a weighted structure on the constant sheaf $\underline{\mathbb{R}^n}$, where the standard inner product is used for the vertex stalks and the weight matrix W_{uv} defines the inner product on edge stalks. A semidefinite matrix defines an inner product on its image in $\mathbb{R}^n$ (or equivalently, the orthogonal complement—with respect to the standard inner product—of its kernel), so it naturally defines a weighted sheaf structure on a sheaf

with vertex stalks $\mathbb{R}^n$ and edge stalks $\operatorname{im} W_{uv}$, with restriction maps the orthogonal projections $\mathbb{R}^n \to \operatorname{im} W_{uv}$ (again with respect to the standard inner product).

A slightly different perspective comes from taking a rank-revealing decomposition $W_{uv} = R_{uv}^* R_{uv}$ of each weight matrix, and set, for each edge $e = u \sim v$, the restriction maps $\mathcal{F}_{u \trianglelefteq e} = \mathcal{F}_{v \trianglelefteq e} = R_{uv}$. This sheaf has vertex stalks $\mathcal{F}(v) = \mathbb{R}^n$ and edge stalks $\mathcal{F}(e) = \mathbb{R}^{\operatorname{rank} W_{uv}}$, with the standard inner products. This construction is unitarily isomorphic to the first; the isomorphism is given by taking the identity on vertex stalks and $R_{uv}|_{\operatorname{im} W_{uv}}$ on edge stalks. A short calculation shows that the sheaf Laplacian of this sheaf is equal to the matrix-weighted Laplacian.

2.6 WEIGHTED SIMPLICIAL COMPLEXES

A recent research program has defined and studied "weighted simplicial complexes" in a sense more general than that used in standard discrete Hodge theory. These results have largely focused on the case where integral domain-valued weights are assigned to simplices [RWW18; Wu+19; Wu+20], but one iteration used a broader definition [Wu+19].

This definition is as follows: Let X be a simplicial complex, R a commutative ring, and M an R-module. A weight function on X is a function $\phi : X \times X \to R$, such that for every simplex $\sigma \in X$ and $x \in M$,

$$\phi(d_i\sigma, d_j d_i \sigma)\phi(\sigma, d_i\sigma)x = \phi(d_j\sigma, d_j d_i \sigma)\phi(\sigma, d_j\sigma)x.$$

There are four simplices involved in this equation, $d_i d_j \sigma \trianglelefteq d_i\sigma, d_j\sigma \trianglelefteq \sigma$. If we think of $\phi(\tau, \sigma)$ as defining a map $\mathcal{F}_{\sigma \trianglelefteq \tau} : M \to M$, we see that this equation is precisely the commutativity condition for these maps to define a cosheaf on X with all stalks equal to M. The paper [Wu+19] goes on to define the cohomology of such weighted complexes, as well as Hodge Laplacians for the case $R = M = \mathbb{R}$.

2.7 UNIQUE GAMES

The unique games conjecture was introduced by Subhash Khot as a question about the computational complexity of computing approximate solutions to certain problems [Kho10]. The goal was to find a universal problem whose complexity of approximation controls that of many other problems. One formulation of the problem that the unique games conjecture considers is the *label cover* problem with unique constraints. An instance of this problem consists of a finite set A, an oriented graph G, and an automorphism φ_e of A associated to each oriented edge of G, so that reversing the orientation of e replaces φ_e with its inverse. (This is precisely the data of a bundle over G with fiber A as for a gain/voltage graph.) The label

cover problem is to find the *satisfiability fraction*—the maximal proportion of edges that can be satisfied by an assignment $x_v \in A$ for each vertex. In bundle-theoretic terms, how close is the bundle to having a section?

If the bundle has a section, it is easy to find, since it must be uniquely determined on each connected component of G by its value at a single vertex. However, finding the satisfiability fraction for a bundle without a section appears to be quite difficult in general. The unique games conjecture makes this precise:

Conjecture (Unique Games Conjecture)**.** *For sufficiently small $\epsilon, \delta > 0$, there exists some constant k such that it is NP-hard to distinguish at least $(1-\delta)$ satisfiable instances of the label cover problem with a label set of size k from at most ϵ-satisfiable instances.*

Recent work [KMS17; KMS18] has proved a relaxed version of this conjecture, the *2-to-2 games conjecture*, which looks for sections of a sheaf of sets over a graph, rather than just a bundle. An instance of the label cover problem with 2-to-2 constraints is a graph G together with a sheaf of sets on G with vertex stalks equal to some set A, and where the fibers $\mathcal{F}^{-1}_{v \trianglelefteq e}(\{x\})$ of the restriction maps have cardinality 2. This means that for a given $x_v \in \mathcal{F}(v)$, there exist two satisfactory extensions x_u to an adjacent vertex u.

The proof of the 2-to-2 games conjecture relies on an interesting construction: the Grassmann graph $\mathrm{Gr}(\mathbb{F}_p^n, k)$ for a vector space over a finite field. This graph has an associated sheaf, analogous to the tautological bundle for a Grassmannian manifold.

Definition 2.7.1 (Grassmann graph)**.** The *Grassmann graph* $\mathrm{Gr}(\mathbb{F}_p^n, k)$ is a graph with one vertex v for each k-dimensional subspace $W_v \leqslant \mathbb{F}_p^n$, and an edge between vertices u and v if $W_u \cap W_v$ is of dimension $k-1$.

There is a tautological cosheaf of $\mathbb{F}_p$-vector spaces over the Grassmann graph, where the stalk $\hat{\mathcal{F}}(v)$ over v is the vector space W_v, the stalk over an edge $e = u \sim v$ is $\hat{\mathcal{F}}(e) = W_u \cap W_v$, and all extension maps are inclusions of subspaces. The sheaf on $\mathrm{Gr}(\mathbb{F}_p^n, k)$ used in the proof of the 2-to-2 games conjecture is a sheaf of sets dual to this cosheaf. The stalk over a vertex v is the set of linear functions $W_v \to \mathbb{F}_p$, with the restriction maps now restriction of functions to a subset of their domain. It turns out that this sheaf of binary functions on the Grassmann graph for $\mathbb{F}_2^n$ can encode a known-hard approximation problem as a subset of a 2-to-2 games problem. A key piece of the proof is showing that 0-cochains of this sheaf which are consistent over most edges do not always have a nearby section; in other words, approximate sections do not always lie close to sections.

2.8 MULTILAYER NETWORKS

Network science has a long history of interesting insights predicated on the notion that connectivity relations within a system are important. These results typically abstract away the nature of the connections, focusing only on whether a given link exists or not. Recently, more attention has been drawn to the fact that in many situations, the nature of these connections is also important. One tool introduced to model these situations is the *multilayer network*.

Definitions vary widely (see [Kiv+14]), but one common formulation is that a multilayer network is represented by an edge-colored multigraph. There is often also some implied structure on the set of vertices, typically defining clusters of vertices that form supernodes, or sets of vertices that all correspond to the same referent. A typical picture—and the origin of the term "multilayer"—is of a series of graphs stacked together over some base space. For instance, in a transportation network, the layers might correspond to different modes of transportation, or in a social network, different layers might correspond to different types of social relationships, e. g., friends, coworkers, acquaintances.

This picture of a stack of graphs fibered over a base graph suggests a model using sheaves. The pushforward of the constant sheaf on the union of the graph layers is one possible sheaf model for multilayer networks. Indeed, the so-called supra-Laplacian of a multilayer network can be seen as the Laplacian of this pushforward sheaf.

It is best not to take this analogy too far. Much of the work done with multilayer networks is difficult to formulate in sheaf-theoretic terms. In many cases the natural semantic interpretations of a sheaf and of a multilayer network are simply different. However, the relationship is interesting to note and may prove fruitful in further interactions between network science and sheaf theory.

2.9 SHEAVES ON GRAPHS

Joel Friedman used a construction he called "sheaves on graphs" to prove the Hanna Neumann conjecture on subgroups of a free group [Fri15]. Friedman's sheaves are in our terminology cellular cosheaves. That is, they consist of an assignment of vector spaces to vertices and edges of a graph, with maps from the spaces over edges to the spaces over their incident vertices. Friedman interprets these as sheaves for a particular Grothendieck topology associated to the graph. The homology of particular cosheaves is related to the properties of certain covering maps of graphs.

The reason for the interest in graph coverings and cosheaves for a problem in group theory is that subgroups of a free group correspond to coverings of a graph

via the fundamental group. Each covering map of graphs $E \to G$ determines a cosheaf on G, the pushforward of the constant sheaf on E. Representing covers in this way allows us to do things like construct kernels and cokernels of maps of covering spaces.

In an offhand remark, Friedman defined Laplacians and adjacency matrices for these sheaves on graphs, equivalent to the Hodge-theoretic definitions, and suggested their study as an extension of spectral graph theory. Subsequent work on sheaves by Friedman and collaborators has focused on homological properties of sheaves associated to graphs, applied to such topics as the graph Riemann-Roch theorem [FF17].

3

GENERALITIES ON SHEAF LAPLACIANS

Having made the appropriate definitions and explored some examples, we can begin to develop more seriously the theory of Laplacians of cellular sheaves. The results here will not involve the spectra of these operators, and so are not properly part of a spectral sheaf theory. However, much of this material will be useful in applications of sheaf Laplacians to engineering and data analysis problems.

3.1 HARMONICITY

The spaces of harmonic cochains $\mathcal{H}(X;\mathcal{F})$ of a sheaf are solutions to the equation $\Delta^k x = 0$. This equation is analogous to Laplace's equation on a manifold without boundary. We might adjust this condition somewhat by asking for the Laplacian to vanish only on a selection of cells, analogous to a boundary value problem in the world of partial differential equations. Given a subcomplex $B \subseteq X$ with complement $U = X \setminus B$, and a cochain $y \in C^k(B;\mathcal{F})$, the degree-k sheaf boundary value problem is

$$\begin{aligned}(\Delta^k_{\mathcal{F}} x)|_U &= 0\\ x|_B &= y\end{aligned}$$

This problem is commonly known as harmonic extension: given a cochain on the boundary B, we wish to extend it to the rest of the complex so that it is harmonic on U.

Solving this problem is a straightforward exercise in linear algebra. Take a block decomposition of $\Delta^k_{\mathcal{F}}$ and write the problem to be solved as

$$\begin{bmatrix}\Delta^k[U,U] & \Delta^k[U,B]\\ \Delta^k[B,U] & \Delta^k[B,B]\end{bmatrix}\begin{bmatrix}x|_U\\ y\end{bmatrix} = \begin{bmatrix}0\\ z\end{bmatrix}.$$

(We will drop the subscript $\mathcal{F}$ for notational convenience.) Since z is a free variable, the only constraint is that $\Delta^k[U,U]x|_U + \Delta^k[U,B]y = 0$. This equation always has a solution when $\operatorname{im}\Delta^k[U,U]$ contains $\operatorname{im}\Delta^k[U,B]$, and this solution is unique if $\Delta^k[U,U]$ is invertible. This problem turns out to be closely related to relative sheaf cohomology.

Proposition 3.1.1. *Let X be a regular cell complex with $B \subseteq X$ a subcomplex, with $\mathcal{F}$ a weighted cellular sheaf on X. Given a k-cochain $x|_B$ defined on B, there exists a k-cochain $x \in C^k(X;\mathcal{F})$ which is the harmonic extension of $x|_B$, and x is unique if and only if the map $H^k(X,B;\mathcal{F}) \to H^k(X;\mathcal{F})$ is zero.*

Proof. First note that $\operatorname{im}\Delta^k[U,B]$ is always contained in $\operatorname{im}\Delta^k[U,U]$. This is a general fact about symmetric positive semidefinite matrices. If $x \in \ker\Delta^k[U,U]$, then its extension by zero to C^k, $\tilde{x}$, is in $\ker\Delta^k$. In particular, this means that $\Delta^k[B,U]x = 0$, and hence x is orthogonal to the image of $\Delta^k[B,U]^* = \Delta^k[U,B]$. Since $\ker\Delta^k[U,U] = (\operatorname{im}\Delta^k[U,U])^\perp$, we see that $\operatorname{im}\Delta^k[U,U] \supseteq \operatorname{im}\Delta^k[U,B]$. Thus, every cochain on B has a harmonic extension to U.

For the question of uniqueness, consider now the direct sum decomposition $C^k(X;\mathcal{F}) = C^k(B;\mathcal{F}) \oplus C^k(X,B;\mathcal{F})$. The coboundary map of $\mathcal{F}$ does not split across this decomposition, but it does have a block structure

$$\delta^k = \begin{bmatrix} \delta^k[B,B] & 0 \\ \delta^k[U,B] & \delta^k[U,U] \end{bmatrix}.$$

The Hodge Laplacian of the complex $C^k(X,B;\mathcal{F})$ is the Hodge Laplacian generated by $\delta[U,U]$, namely

$$\Delta^k(X,B) = (\delta^k[U,U])^*\delta^k[U,U] + \delta^{k-1}[U,U](\delta^{k-1}[U,U])^*.$$

Meanwhile, the $[U,U]$ block of the matrix Δ^k is

$$(\delta^k[U,U])^*\delta^k[U,U] + \delta^{k-1}[U,U](\delta^{k-1}[U,U])^* + \delta^{k-1}[U,B](\delta^{k-1}[U,B])^* \\ = \Delta^k(X,B) + \delta^{k-1}[U,B](\delta^{k-1}[U,B])^*.$$

We now need to determine when $\Delta^k[U,U]$ is invertible. Since $\Delta^k[U,U] = \Delta^k(X,B) + \delta^{k-1}[U,B](\delta^{k-1}[U,B])^*$, we know that $\ker\Delta^k[U,U] \subseteq \mathcal{H}^k(X,B;\mathcal{F})$. Indeed, it is equal to $\mathcal{H}^k(X,B;\mathcal{F}) \cap (\operatorname{im}\delta^{k-1}[U,B])^\perp$. On Hodge representatives for cohomology, the connecting map $d : \mathcal{H}^{k-1}(B;\mathcal{F}) \to \mathcal{H}^k(X,B;\mathcal{F})$ is computed by $\delta^{k-1}[U,B]$ restricted to harmonic cochains, so we have $\ker\Delta^k[U,U] = (\operatorname{im} d)^\perp$, and hence this is zero if and only if $\operatorname{im} d = \mathcal{H}^k(X,B;\mathcal{F})$. By exactness, this condition is equivalent to the vanishing of the map $H^k(X,B;\mathcal{F}) \to H^k(X;\mathcal{F})$ in the relative cohomology long exact sequence. □

The image of the relative Hodge cohomology $\mathcal{H}(X,B;\mathcal{F})$ in $\mathcal{H}(X;\mathcal{F})$ is essentially the "harmonic extension of zero." It therefore controls the indeterminacy of har-

monic extensions of nonzero boundary conditions. Of course, we can also solve the harmonic extension problems for the up- and down-Laplacians alone; harmonic extension for the up-Laplacian is related to effective resistance, as we will see in Section 3.2.

Harmonic extension has another interpretation in terms of a variational problem. Recall the Laplacian quadratic form $\mathcal{E}^k(x) = \langle x, \Delta^k_{\mathcal{F}} x\rangle$. We may think of this as giving the energy associated to each cochain of $\mathcal{F}$. The harmonic cochains $\mathcal{H}^k$ are those cochains x with $\mathcal{E}^k(x) = 0$. We can thus think of harmonic extension as a minimization of $\mathcal{E}^k$ subject to boundary conditions:

$$\begin{aligned} &\min \langle x, \Delta^k_{\mathcal{F}} x\rangle \\ &\text{s.t. } \Pi_B x = y \end{aligned}$$

where Π_B is the orthogonal projection $C^k(X;\mathcal{F}) \to C^k(B;\mathcal{F})$. This problem is equivalent to the formulation in terms of the vanishing of $\Delta^k x$ on $X \setminus B$. To see this, note that this is a convex program with linear constraints; the Lagrangian is

$$\mathcal{L}(x,\lambda) = \langle x, \Delta^k \mathcal{F} x\rangle + \langle \lambda, P_B x - y\rangle$$

and the Lagrange multiplier conditions are

$$\begin{aligned} 2\Delta^k_{\mathcal{F}} x &= \Pi^*_B \lambda \\ \Pi_B x &= y. \end{aligned}$$

$\Pi^*_B \lambda$ is a k-cochain which is zero on stalks over cells not in B, and has values determined by λ on stalks over cells in B. Thus the optimality condition is equivalent to

$$\begin{aligned} (\Delta^k_{\mathcal{F}} x)|_U &= 0 \\ x|_B &= y. \end{aligned}$$

Casting harmonic extension as an optimization problem gives us another way to see that every boundary condition has a harmonic extension, since the optimization problem here is always feasible.

3.1.1 $O(n)$-*bundles and the maximum modulus principle*

Harmonic functions on subsets of $\mathbb{R}^n$ have many nice properties. Among these is the maximum principle: A harmonic function $D \to \mathbb{R}$ attaining a maximum or minimum on the interior of D must be constant. Similarly, a holomorphic function attaining its maximum modulus on the interior of its domain is constant. Such a

result does not hold in general for harmonic cochains on sheaves, but something analogous does work for $O(n)$-bundles.

Lemma 3.1.1. *Let $\mathcal{F}$ be an $O(n)$-bundle on a cell complex X, with constant vertex weights $\alpha_v = 1$ and arbitrary edge weights α_e (as discussed in Section 1.10). If $x \in C^0(X;\mathcal{F})$ is harmonic at a vertex v, then*

$$x_v = \frac{1}{d_v} \sum_{\substack{u,v \trianglelefteq e \\ u \neq v}} \mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{u \trianglelefteq e} x_u,$$

where $d_v = \sum_{v \trianglelefteq e} \|\mathcal{F}_{v \trianglelefteq e}\|^2 = \sum_{v \trianglelefteq e} \alpha_e^2$.

Proof. This is a simple consequence of the block structure of the sheaf Laplacian. If x is harmonic at v, $(L_{\mathcal{F}} x)_v = 0$. Evaluating the Laplacian at x, we get

$$0 = \sum_{v \trianglelefteq e} \mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{v \trianglelefteq e} x_v - \sum_{\substack{u,v \trianglelefteq e \\ u \neq v}} \mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{u \trianglelefteq e} x_u.$$

Because $\mathcal{F}$ is an $O(n)$-bundle, we have $\mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{v \trianglelefteq e} = \sum_{v \trianglelefteq e} \|\mathcal{F}_{v \trianglelefteq e}\|^2 \operatorname{id}_{\mathcal{F}(v)} = d_v \operatorname{id}_{\mathcal{F}(v)}$. Rearranging terms yields the formula. □

Theorem 3.1.1 (Maximum modulus principle)**.** *Let $\mathcal{F}$ be an $O(n)$-bundle on a cell complex X, with $\alpha_v = 1$ for all 0-cells v. Suppose B is a* thin *subcomplex of X, i.e.one such that $X \setminus \mathrm{st}(B)$ is connected, and every 0-cell of B is incident to a cell not in B. If $x \in C^0(X;\mathcal{F})$ is harmonic on $X \setminus B$ and attains its maximum stalkwise norm at a vertex in $X \setminus B$, then it has constant stalkwise norm.*

Proof. Suppose that $v \in X \setminus B$ is a vertex for which $\|x_v\| \geqslant \|x_u\|$ for all 0-cells $u \in X$. Then we have (using the lemma):

$$\begin{aligned} \|x_v\| = \frac{1}{d_v} \Big\| \sum_{\substack{u,v \trianglelefteq e \\ u \neq v}} \mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{u \trianglelefteq e} x_u \Big\| &\leqslant \frac{1}{d_v} \sum_{\substack{u,v \trianglelefteq e \\ u \neq v}} \|\mathcal{F}^*_{v \trianglelefteq e} \mathcal{F}_{u \trianglelefteq e} x_u\| \\ &= \frac{1}{d_v} \sum_{\substack{u,v \trianglelefteq e \\ u \neq v}} \alpha_e^2 \|x_u\| \leqslant \frac{1}{d_v} \sum_{v \trianglelefteq_1 e} \alpha_e^2 \|x_v\| = \|x_v\|. \end{aligned}$$

Equality therefore holds throughout, which, because $\|x_v\|$ is maximal, forces $\|x_v\| = \|x_u\|$ for all neighbors u of v. Since $X \setminus \mathrm{st}(B)$ is connected, this local constancy now propagates across all vertices of X, including those of B, since every vertex of B is adjacent to one not in B. Thus x has constant stalkwise norm. □

Corollary. *Let* B *be a thin subcomplex of* X*, with* $\mathcal{F}$ *an* $O(n)$*-bundle on* X *as before. If* $x \in C^0(X;\mathcal{F})$ *is harmonic on* $X \setminus B$*, then it attains its maximum stalkwise norm on* B*.*

We can weaken the thinness condition somewhat while preserving the truth of the corollary. For instance, we might only require that $X \setminus \mathrm{st}(B)$ be connected. Particularizing to the constant sheaf (which is an $O(1)$-bundle), we get a slightly weaker version of a well-known result discussed in [Suno8], which does not require taking norms. One might also view this result as generalizing the fact that sections of $O(n)$-bundles have constant norm, since this corresponds to the choice $B = \varnothing$.

3.2 EFFECTIVE RESISTANCE

The relationship between graph Laplacians and the analysis of electrical networks is well established. If we consider a weighted graph G as a circuit of resistors with conductivities given by the edge weights, the Laplacian of G becomes a map taking node voltages to net current flow at each node. That is, if x is a vector of voltages at nodes of G, $L_G x$ is a vector representing the induced net current flowing into or out of each node. The Laplacian quadratic form $\langle x, L_G x\rangle$ measures the power dissipated in the circuit for a given voltage distribution x. Conversely, the (Moore-Penrose) pseudoinverse $L_G^\dagger$ is a map from current flows to voltage distributions. (There is an indeterminacy here; by definition, the pseudoinverse will choose the voltage distribution which averages to zero.)

We can break this relationship down into finer pieces. Consider the dual vector spaces $C^k(G;\mathbb{R})$ and $C_k(G;\mathbb{R})$. Elements of C^k represent voltages or potential differences; elements of C_k represent currents. The coboundary map $\delta : C^0(G;\mathbb{R}) \to C^1(G;\mathbb{R})$ takes a voltage distribution on nodes and computes the potential difference over each edge of G. The boundary map $\partial : C_1(G;\mathbb{R}) \to C_0(G;\mathbb{R})$ takes currents on edges and calculates the net current flow into or out of each vertex. The natural pairing $C^1 \otimes C_1 \to \mathbb{R}$ calculates power dissipation: it multiplies current by voltage on each edge. Ohm's law gives an isomorphism $R : C_1(G;\mathbb{R}) \to C^1(G;\mathbb{R})$, saying that current on each edge is proportional to voltage over each edge, with the resistance being the constant of proportionality. This isomorphism induces an inner product on C^1 and C_1, giving the familiar formulas $P = I^2R = V^2/R$. The Laplacian of G may then be seen as the map $L = \partial \circ R^{-1} \circ \delta$ from $C^0(G;\mathbb{R})$ to $C_0(G;\mathbb{R})$. The pairing $\langle x, L_G x\rangle$ is then equal to $\langle \delta x, \delta x\rangle$, which sums V^2/R for each edge, computing power dissipation.

Kirchoff's laws imply that if we fix certain nodes of a circuit to given voltages and leave the others free, the net current flow on the free nodes is zero; that is, $L_G x = 0$ on internal nodes. This is precisely the equation for harmonic extension on a graph. By the dual formulation in terms of energy minimization, a circuit

with voltage boundary conditions assumes the current flow that minimizes total power dissipation. The same holds for current boundary conditions.

A common tool in circuit analysis is to reduce the relationship between two nodes to an equivalent circuit. In the case of resistive networks, one may replace the entire network by a single resistor between the two nodes of interest. The resistance of this resistor is the *effective resistance* between the two nodes.

We can calculate the effective resistance between two circuit nodes by using either voltage boundary conditions or current boundary conditions. The pseudoinverse of L_G computes the map from current inputs and outputs to voltage distributions. We choose two vertices u and v, and let y be a current distribution with $y_u = -1$, $y_v = 1$, and $y_w = 0$ for any other vertex. Then $L_G^\dagger y$ is a voltage distribution on the nodes of G. In particular, the potential difference between u and v is $(L_G^\dagger y)_v - (L_G^\dagger y)_u = \langle y, L_G^\dagger y\rangle$. Since the current flowing between u and v is 1, this voltage is equal to the effective resistance between the two nodes, i. e.the constant of proportionality between current and voltage.

Conversely, we can set a boundary condition $x|_B$ on $B = \{u, v\}$ with $x_u = 0$ and $x_v = 1$, and extend it to a cochain x harmonic on $G \setminus B$. This amounts to finding the voltage distribution that satisfies Kirchoff's voltage law on U. Then $L_G x$ is zero outside of B, and at u and v it contains a current flow. Thus, $\langle x, L_G x\rangle$ is the net current flow into v, and since the voltage difference between u and v is 1, this is equal to the conductance (reciprocal of the resistance) between the two nodes.

These two dual calculations bifurcate further when we extend them to sheaves. A generalization of the current boundary conditions gives us an initial definition of effective resistance for cosheaves. Of course, this definition can be extended to sheaves by duality.

Definition 3.2.1 (Effective resistance)**.** Let $\hat{\mathcal{F}}$ be a weighted cosheaf on a complex X, and $a, b \in \ker \partial_k = Z_k(X; \hat{\mathcal{F}})$ homologous k-cycles. The effective resistance between a and b is

$$R_{\text{eff}}(a, b) = \min_{c \in C_{k+1}(X;\hat{\mathcal{F}})} \|c\| \quad \text{s.t. } \partial c = b - a.$$

One may think of this as the ℓ^2 size of the minimal witness that a and b are in fact homologous.

Proposition 3.2.1. *For any homologous* $a, b \in Z_k(X; \hat{\mathcal{F}})$, $R_{\mathit{eff}}(a, b) = \langle b - a, (\Delta_+^k)_{\hat{\mathcal{F}}}^\dagger (b - a)\rangle$.

Proof. Since $(\Delta_+^k)_{\hat{\mathcal{F}}} = \partial_{k+1}\partial_{k+1}^*$, a standard fact about the Moore-Penrose pseudoinverse implies that $(\Delta_+^k)_{\hat{\mathcal{F}}}^\dagger = (\partial_{k+1}^\dagger)^* \partial_{k+1}^\dagger$. Therefore,

$$\langle b - a, (\Delta_+^k)_{\hat{\mathcal{F}}}^\dagger (b - a)\rangle = \langle \partial_{k+1}^\dagger (b - a), \partial_{k+1}^\dagger (b - a)\rangle = \|\partial_{k+1}^\dagger (b - a)\|^2.$$

By the least squares property of the pseudoinverse, if $b - a \in \operatorname{im} \partial_{k+1}$, $\partial_{k+1}^{\dagger}(b-a)$ is the c of smallest norm such that $\partial_{k+1} c = b - a$. □

The characteristic case is when a and b are each supported on distinct cells. In dimension zero, we think of a as being an $\hat{\mathcal{F}}$-valued current flow into a vertex and b as its corresponding outwards flow. The effective resistance is a measure of the total amount of current flow (as measured by the ℓ^2 norm on each edge) in the graph induced by this input-output flow. This is why we require a and b to be homologous: current must be conserved.

When $\hat{\mathcal{F}}$ is a weighted constant cosheaf $\overline{\mathbb{R}}$ on a graph, this reduces exactly to effective resistance for an electrical network. This is because a 0-chain supported on a vertex v has a canonical corresponding 0-chain supported on any other vertex u, so to compute the effective resistance between u and v we consider indicator cochains supported on u and v. For a general cosheaf, we need to allow for nonconstant and nonunique homologous extensions. However, in the case of a matrix-weighted graph, we again have a canonical pairing of 0-chains supported on individual vertices. As a result, we can produce an analogue of the effective resistance between vertices for such a cosheaf, letting $R_{\text{eff}}(u,v) = L^{\dagger}[u,u] + L^{\dagger}[u,v] - 2L^{\dagger}[u,v]$. This is the matrix representing the quadratic form $Q(x) = R_{\text{eff}}(x_u, x_v)$, where x_u and x_v have the same value over different vertices.

The cosheaf effective resistance as defined only depends on the difference $b - a$ of the cochains, so it is really a quadratic form on $Z_k(X;\hat{\mathcal{F}})$, represented by the matrix $(\Delta_+^k)^{\dagger}_{\hat{\mathcal{F}}}$. Restricting to chains supported on a subcomplex amounts to taking a principal submatrix of $(\Delta_+^k)^{\dagger}_{\hat{\mathcal{F}}}$. In particular, in dimension zero, if the support is a pair of vertices, we may think of $(\Delta_+^0)^{\dagger}_{\hat{\mathcal{F}}}[\{u,v\},\{u,v\}]$ as describing the effective resistance between the vertices u and v. This matrix represents a quadratic form calculating the power dissipation if a given current enters one vertex and leaves the other.

The calculation of effective conductance using voltage boundary conditions also generalizes to sheaves. However, we need to choose a particular boundary subcomplex.

Definition 3.2.2. Let $\mathcal{F}$ be a sheaf on a complex X, with $B \subseteq X$ a subcomplex. The *effective conductance* of a cochain $x_B \in C^k(B;\mathcal{F})$ is

$$\min_{x \in C^k(X;\mathcal{F})} \|\delta^k x\|^2$$
$$\text{s.t. } x|_B = x_B$$

We denote the effective conductance $G_{\text{eff}}(x_B)$.

This is precisely the optimization problem for harmonic extension, with the full Laplacian replaced with the up-Laplacian. If we restrict to $x_B \in (B^{k-1}(X;\mathcal{F}))^\perp$, these two problems are equivalent. (This restriction corresponds to our restriction to $Z_k(X;\hat{\mathcal{F}})$ for cosheaf effective resistance.) Effective conductance is a quadratic form on $C^k(B;\mathcal{F})$. It contains the same information as the effective resistance quadratic form on $C_k(B;\mathcal{F}^*)$.

Proposition 3.2.2. *Let $\mathcal{F}$ be a sheaf on a complex X, with $\mathcal{F}^*$ its dual cosheaf, and $B \subseteq X$ a boundary subcomplex. The matrix representing the effective conductance quadratic form on $B^k(B;\mathcal{F})^\perp$ is the (pseudo)inverse of the matrix representing the effective resistance quadratic form on $Z_k(B;\mathcal{F}^*)$.*

Proof. Note that $B^k(B;\mathcal{F})^\perp \simeq Z_k(B;\mathcal{F}^*)$ since the two chain complexes are adjoints of each other. For $x_B \in B^k(B;\mathcal{F})^\perp$, the harmonic extension to U is

$$x_U = -(\Delta^k_+)_{\mathcal{F}}[U,U]^\dagger(\Delta^k_+)_{\mathcal{F}}[U,B]x_B,$$

and hence its norm is

$$\langle x_B, (\Delta^k_+)_{\mathcal{F}}[B,B]x_B\rangle - \langle (\Delta^k_+)_{\mathcal{F}}[U,B]x_B, (\Delta^k_+)_{\mathcal{F}}[U,U]^\dagger(\Delta^k_+)_{\mathcal{F}}[U,B]x_B\rangle.$$

Thus the matrix representing effective conductance on B is

$$G_{\text{eff}}(B) = (\Delta^k_+)_{\mathcal{F}}[B,B] - (\Delta^k_+)_{\mathcal{F}}[B,U](\Delta^k_+)_{\mathcal{F}}[U,U]^\dagger(\Delta^k_+)_{\mathcal{F}}[U,B],$$

the Schur complement of $(\Delta^k_+)_{\mathcal{F}}[U,U]$. Conversely, the effective resistance matrix is $(\Delta^k_+)^\dagger_{\mathcal{F}}[U,U]$. But a submatrix of the pseudoinverse is the inverse of the corresponding Schur complement. □

A particularly nice representation exists for the effective resistance matrix corresponding to the boundary of a single $(k+1)$-cell σ. A canonical subspace of k-cycles supported on this boundary is $\operatorname{im}\partial|_\sigma$. This allows us to construct a matrix which we call the effective resistance of σ:

$$R_{\text{eff}}(\sigma) = \partial^*_\sigma(\Delta^k_+)^\dagger_{\hat{\mathcal{F}}}\partial_\sigma.$$

This defines a quadratic form on $\hat{\mathcal{F}}(\sigma)$, which represents the importance of different subspaces of $\hat{\mathcal{F}}(\sigma)$ in terms of generating current flows on $\partial\sigma$. If $x \in \hat{\mathcal{F}}(\sigma)$ generates a current flow on $\partial\sigma$ which cannot be produced by any other chain, then $\langle x, R_{\text{eff}}(\sigma)x\rangle = \langle x, x\rangle$. Conversely, if some well-dispersed chain y also has $\partial y = \partial x$, then $\langle x, R_{\text{eff}}(\sigma)x\rangle \leqslant \langle y, y\rangle \ll \langle x, x\rangle$. This property of effective resistance plays an important role in enabling sparsification algorithms for graphs, complexes, and sheaves, as will be discussed in Chapter 5.

3.3 IDENTIFYING SHEAF LAPLACIANS

Laplacians of weighted graphs are simple to identify: they are the symmetric matrices with nonnegative diagonal entries, nonpositive off-diagonal entries, and row (or column) sums zero. This characterization determines the set of graph Laplacians as a convex cone in the space of symmetric matrices. A similar characterization for Laplacians of sheaves on graphs requires deeper investigation. Consider the case of sheaves with fixed-dimensional vertex stalks and edge stalks of arbitrary dimension. (This indeterminacy of edge stalks is reasonable since, as discussed in Section 1.7.1, we cannot identify edge stalk dimensions from a sheaf Laplacian.) The coboundary matrix of such a sheaf has a block structure determined by the decomposition of C^0 and C^1 into direct sums of stalks. As a result of the underlying graph structure, every block row of δ has at most two nonzero blocks. In this situation, we say that δ is 2-block row sparse.

The question of the structure of sheaf Laplacians then boils down to the question of the properties of matrices $\delta^*\delta$, where δ is 2-block row sparse. When vertex stalks are of dimension 1, this was studied by Boman et al. [Bom+05], who called such matrices *factor width two*. Their main result is that a matrix L has factor width two if and only if it is *generalized diagonally dominant*: there exists some positive diagonal matrix D such that

$$(DLD)_{ii} \geqslant \sum_{i \neq j} \left|(DLD)_{ij}\right|$$

for all i. One direction of the proof is a special case of Proposition 3.3.1, while the other direction involves showing that all factor-width two matrices are symmetric positive definite H-matrices. An H-matrix is a matrix A whose *comparison matrix* $M(A)$, given by changing signs so that all diagonal elements are positive and all off-diagonal elements are negative, is an M-matrix. A matrix A is an M-matrix if $A = sI - B$ for some nonnegative matrix B and s larger than the spectral radius of B.

Definition 3.3.1. Let L be a symmetric block matrix. We say that L has *block factor width two* if there exists a matrix B with the same block structure as L on its columns which is 2-block row sparse, and $L = B^T B$.

The analogous notion of diagonal dominance takes a little more to develop. Recall that the *absolute value* of a matrix A is the symmetric positive semidefinite matrix $|A| = \sqrt{AA^T}$. This construction is useful in connection with the *polar decomposition* $A = |A|U$, where U is a unitary matrix.

Definition 3.3.2. A symmetric block matrix L is *block diagonally dominant* if

$$L_{ii} - \sum_{j \neq i} \left|L_{ij}\right|$$

is positive semidefinite for all i. (In particular, note that this requires that all diagonal blocks be positive semidefinite.)

Proposition 3.3.1. *If L is symmetric and block diagonally dominant then it has a factorization of block width two, i. e. $L = B^T B$ where the block rows of B have at most two nonzero blocks.*

Proof. Write $L_{ij} = U_{ij}\Sigma_{ij}V_{ij}^T$ for $i < j$ via the singular value decomposition. We use the compact form of the SVD, where Σ is a square diagonal matrix and U and V have the same number of columns. Then $\left|L_{ij}\right| = U_{ij}\Sigma_{ij}U_{ij}^T$ for $i < j$ and $\left|L_{ij}\right| = V_{ij}\Sigma_{ij}V_{ij}^T$ for $i > j$. Let

$$C_{ij} = \begin{bmatrix} \cdots & \Sigma_{ij}^{1/2}U_{ij}^T & \cdots & \Sigma_{ij}^{1/2}V_{ij}^T & \cdots \end{bmatrix}$$

for $i < j$. Note that

$$C_{ij}^T C_{ij} = \begin{bmatrix} & \vdots & & \vdots & \\ \cdots & U_{ij}\Sigma_{ij}U_{ij}^T & \cdots & U_{ij}\Sigma_{ij}V_{ij}^T & \cdots \\ & \vdots & & \vdots & \\ \cdots & V_{ij}\Sigma_{ij}U_{ij}^T & \cdots & V_{ij}\Sigma_{ij}V_{ij}^T & \cdots \\ & \vdots & & \vdots & \end{bmatrix} = \begin{bmatrix} & \vdots & & \vdots & \\ \cdots & \left|L_{ij}\right| & \cdots & L_{ij} & \cdots \\ & \vdots & & \vdots & \\ \cdots & L_{ji} & \cdots & \left|L_{ji}\right| & \cdots \\ & \vdots & & \vdots & \end{bmatrix}.$$

Let $D_i = \begin{bmatrix} \cdots & \sqrt{L_{ii} - \sum_{i \neq j} \left|L_{ij}\right|} & \cdots \end{bmatrix}$, noting that this exists by block diagonal dominance. Then

$$D_i D_i^T = \begin{bmatrix} & \vdots & \\ \cdots & L_{ii} - \sum_{i \neq j} \left|L_{ij}\right| & \cdots \\ & \vdots & \end{bmatrix}.$$

We then let the matrix B be the vertical concatenation of the matrices C_{ij} and the matrices D_i. Then a quick computation shows that $B^T B = \sum_{i<j} C_{ij}^T C_{ij} + \sum_i D_i^T D_i = L$. □

We will say that L is *scaled block diagonally dominant* if there exists an invertible block diagonal matrix D such that $D^T L D$ is block diagonally dominant. We therefore have the following corollary:

Corollary. *If L is symmetric and scaled block diagonally dominant, then it has a factorization of block width two.*

Proof. If $DLD^T = B^TB$ is a factorization of block width two, $L = (BD^{-1})^T BD^{-1}$ is a factorization of L of block width two. □

When we interpret L as a degree-0 sheaf Laplacian, the scaling D corresponds to a change of basis and inner product for vertex stalks.

A similar but more restrictive definition was made by Kyng et al. [Kyn+16] as a generalization of connection Laplacians. Their definition of block diagonal dominance requires that

$$\|L_{ii}\| \geqslant \sum_{i \neq j} \|L_{ij}\|$$

for all i. The authors showed that all such matrices could be written as a sum $B^TB + C$, where B is a 'unitary vertex-edge transfer matrix' (the coboundary matrix of a weighted $U(n)$-bundle on a multigraph) and C is a block diagonal matrix with positive semidefinite blocks. We will call such matrices *block norm diagonally dominant*. A simple corollary of this result is then that scaled block norm diagonally dominant matrices are sheaf Laplacians.

However, there are sheaf Laplacians which are not block norm diagonally dominant. Consider, for instance, Laplacians of matrix weighted graphs with weights that are not full rank. Unless the kernels of the weights align in very precise ways, the sum of norms of weights will be greater than the norm of the sum of weights.

Take as a concrete example the sheaf in Figure 3.1, where vertex stalks are $\mathbb{R}^2$, edge stalks are $\mathbb{R}$, and the restriction maps ρ_i are given by the matrices

$$\rho_1 = \begin{bmatrix} 2 & 0 \end{bmatrix}, \quad \rho_2 = \begin{bmatrix} \sqrt{3} & -1 \end{bmatrix}, \quad \rho_3 = \begin{bmatrix} \sqrt{3} & 1 \end{bmatrix}.$$

The Laplacian of this sheaf is

$$L_{\mathcal{F}} = \left[\begin{array}{cc|cc|cc|cc} 6 & 0 & 0 & 0 & -3 & -\sqrt{3} & -3 & \sqrt{3} \\ 0 & 6 & 0 & -4 & -\sqrt{3} & -1 & \sqrt{3} & -1 \\ \hline 0 & 0 & 6 & 0 & -3 & \sqrt{3} & -3 & -\sqrt{3} \\ 0 & -4 & 0 & 6 & \sqrt{3} & -1 & -\sqrt{3} & -1 \\ \hline -3 & -\sqrt{3} & -3 & \sqrt{3} & 6 & 0 & 0 & 0 \\ -\sqrt{3} & -1 & \sqrt{3} & -1 & 0 & 6 & 0 & -4 \\ \hline -3 & \sqrt{3} & -3 & -\sqrt{3} & 0 & 0 & 6 & 0 \\ \sqrt{3} & -1 & -\sqrt{3} & -1 & 0 & -4 & 0 & 6 \end{array}\right].$$

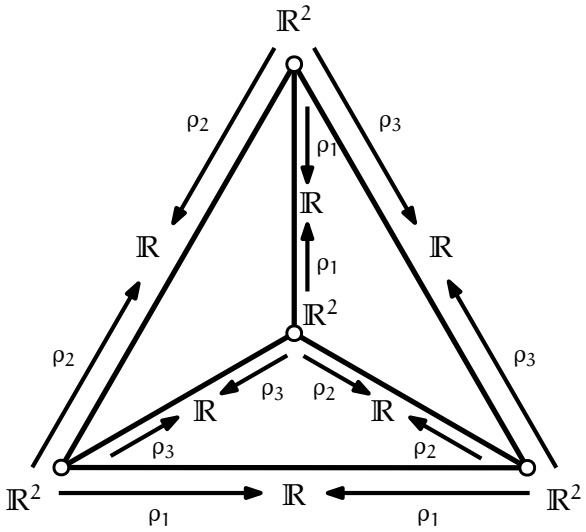


Figure 3.1: A sheaf whose Laplacian is not block norm diagonally dominant.

The norms of the 2×2 off-diagonal blocks on each row are all 4, for a total sum of 12, but the norm of the corresponding diagonal block is 6. On the other hand, this matrix has block row sums 0, so it is clearly block diagonally dominant in our weaker sense.

3.3.1 *Identifying Laplacians for other types of sheaves*

Laplacians of $O(n)$-bundles and matrix-weighted graphs are much simpler to identify than arbitrary sheaf Laplacians. This is in part because we assume a standard basis for vertex stalks, which restricts the properties of the corresponding matrices further.

Proposition 3.3.2. *A symmetric matrix* L *with* $n \times n$ *blocks is the degree-0 Laplacian of an* $O(n)$*-bundle on a simple graph with constant vertex weights* α_v *if and only if*

1. *its diagonal blocks are scalar multiples of the identity matrix*
2. *its nonzero off-diagonal blocks are scalar multiples of orthogonal matrices, and*
3. *for each block row* i, $\|L_{ii}\| = \sum_{j \neq i} \|L_{ij}\|$.

Proof. Suppose L is the degree-0 Laplacian of an $O(n)$-bundle $\mathcal{F}$. Then $L_{ii} = \sum_{i \trianglelefteq e} \mathcal{F}^*_{i \trianglelefteq e} \mathcal{F}_{i \trianglelefteq e} = \sum_{i \trianglelefteq e} \alpha_e^2 I$, and $L_{ij} = \mathcal{F}^*_{i \trianglelefteq e} \mathcal{F}_{j \trianglelefteq e} = \alpha_e^2 \rho_{ij}$, where e is the edge

between vertices i and j and $\rho_{ij} \in O(n)$. These blocks then clearly satisfy conditions 1–3.

Conversely, if L is a block matrix satisfying these conditions, let G be the graph determined by the block sparsity pattern of L. We will construct an $O(n)$-bundle $\mathcal{F}$ on G whose Laplacian is L. If e is the edge between i and j, with $i < j$, let $\mathcal{F}_{i \trianglelefteq e} = \|L_{ij}\|^{1/2} I$ and $\mathcal{F}_{j \trianglelefteq e} = \|L_{ij}\|^{-1/2} L_{ij}$. Then $\mathcal{F}$ is an $O(n)$-bundle with edge weights $\alpha_e = \|L_{ij}\|^{1/2}$ and $(L_{\mathcal{F}})_{ij} = L_{ij}$. Further, $(L_{\mathcal{F}})_{ii} = \sum_{j \neq i} \|L_{ij}\| I$, which is equal to L_{ii} by properties 2 and 3. □

Allowing vertex weights to vary amounts to relaxing this condition to a scaled version: there must be some diagonal matrix D whose blocks are scalar multiples of the identity such that DLD satisfies the conditions for Proposition 3.3.2.

The characterization of Laplacians of matrix-weighted graphs is even simpler, and closer to the characterization of weighted graph Laplacians.

Proposition 3.3.3. *A symmetric matrix L with $n \times n$ blocks is the Laplacian of a matrix-weighted graph if and only if*

1. *its diagonal blocks are positive semidefinite matrices*
2. *its off-diagonal blocks are negative semidefinite matrices, and*
3. *for each block row i, $\sum_j L_{ij} = 0$.*

These conditions are almost automatic: letting $W_{ij} = -L_{ij}$ gives the structure of a matrix-weighted graph.

3.3.2 *Cones of sheaf Laplacians*

Once we pick dimensions for vertex stalks, the set of matrices which arise as degree-0 sheaf Laplacians forms a convex cone in the space of symmetric matrices. From the definition $L_{\mathcal{F}} = \delta_{\mathcal{F}}^* \delta$, we can write

$$L_{\mathcal{F}} = \sum_e \delta_e^* \delta_e := \sum_e L_e,$$

where δ_e is the map $C^0(X; \mathcal{F}) \to \mathcal{F}(e)$ given by the corresponding row of $\delta_{\mathcal{F}}$. Since we have only fixed dimensions of vertex stalks, we may take $\dim \mathcal{F}(e)$ as large as we want. In particular, for an edge e between vertices u and v, if $\dim \mathcal{F}(e) \geqslant \dim \mathcal{F}(u) + \dim \mathcal{F}(v)$, L_e can be any positive semidefinite matrix with the appropriate sparsity pattern. (That is, $L_e[a, b] \neq 0$ if and only if $a, b \in \{u, v\}$.) If $L_{\mathcal{F}}$ and $L_{\mathcal{G}}$ are sheaf Laplacians with the same vertex stalk dimensions, we have

$$L_{\mathcal{F}} + L_{\mathcal{G}} = \sum_e (L_e)_{\mathcal{F}} + (L_e)_{\mathcal{G}} = \sum_e (L_e)_{\mathcal{F}+\mathcal{G}}.$$

Therefore, the sum of two sheaf Laplacians is a sheaf Laplacian, and a positive scalar multiple of a sheaf Laplacian is also a sheaf Laplacian, so sheaf Laplacians form a convex cone.

Laplacians of matrix-weighted graphs also form a convex cone. This is a straightforward consequence of their characterization in Proposition 3.3.3. A positive scalar multiple of a matrix-weighted Laplacian clearly satisfies all three conditions. Since the sum of positive semidefinite matrices is positive semidefinite, conditions 1 and 2 are satisfied for positive linear combinations of matrix-weighted Laplacians. But condition 3 is a linear condition, and hence is satisfied by any linear combination of matrix-weighted Laplacians.

Connection Laplacians (those corresponding to $O(n)$-bundles) do not form a convex cone. This is easy to see because in general the sum of two orthogonal matrices is not an orthogonal matrix. However, they generate a proper subcone of the cone of sheaf Laplacians. To characterize this cone precisely, we will define an auxiliary notion associated with a block norm diagonally dominant matrix L.

Definition 3.3.3. Let L be a block norm diagonally dominant matrix with $n \times n$ blocks. The *degree vector* of L is the vector d with entries $d_i = \|L_{ii}\|$. The *excess degree vector* of L is the vector d^e with entries $d^e_i = d_i - \sum_{i \neq j} \|L_{ij}\|$.

Proposition 3.3.4. *A matrix L with $n \times n$ blocks is a positive linear combination of degree-0 Laplacians of $O(n)$-bundles if and only if the following conditions are satisfied:*

1. *L is block norm diagonally dominant.*
2. *The diagonal blocks of L are positive scalar multiples of the identity.*
3. *The excess degree vector d^e of L is a positive linear combination of indicator vectors on sets of cardinality 2.*

Proof. The "only if" direction is straightforward, since all three conditions are satisfied by connection Laplacians and preserved under positive linear combinations. The only point of difficulty is to see that condition 3 is preserved under positive linear combinations. But this is straightforwardly true for combinations of connection Laplacians on graphs with a single edge, and every connection Laplacian is a sum of connection Laplacians corresponding to a graph with a single edge.

For the reverse direction, we first observe that every real $n \times n$ matrix A of norm at most 1 is a convex combination of orthogonal matrices.[1] By the singular value decomposition $A = U\Sigma V^T$ we only need to show that this holds for every diagonal matrix with entries in $[0, 1]$. It suffices to show, then, that every vector in $[0, 1]^n$ is a

1 This is in fact a special case of the Russo-Dye theorem [Gar84], which states that the closed convex hull of the unitary elements of a C^* algebra is equal to the closed unit ball.

convex combination of vectors in $\{-1, 1\}^n$. But this is clearly true, since these points form the vertices of a cube in $\mathbb{R}^n$ containing $[0, 1]^n$.

Now, suppose L satisfies conditions 1-3. We will first construct a Laplacian matrix $\tilde{L} = \sum_e L_e$ such that $L - \tilde{L}$ is block diagonal. For each edge $e = i \sim j$, we construct the matrix

$$L_e = \begin{bmatrix} \|L_{ij}\| I & L_{ij} \\ L_{ji} & \|L_{ij}\| I \end{bmatrix}.$$

Since $\|L_{ij}\|^{-1} L_{ij}$ has norm 1, it is a convex combination $\sum_k \lambda_k^{ij} \rho_k^{ij}$ where ρ_k^{ij} is an orthogonal matrix for each k. We can then write

$$L_e = \|L_{ij}\| \sum_k \begin{bmatrix} \lambda_k^{ij} I & \rho_k^{ij} \\ (\rho_k^{ij})^T & \lambda_k^{ij} \end{bmatrix}.$$

Extending each matrix L_e to the size of L by placing the nonzero blocks in the positions corresponding to the edge e, we let $\tilde{L} = \sum_e L_e$. Note that $\tilde{L}$ satisfies $\|\tilde{L}_{ii}\| = \sum_{i \neq j} \|L_{ij}\|$, and that $L - \tilde{L}$ is zero except on diagonal blocks, which are nonnegative scalar multiples of I. The vector of norms of diagonal blocks of $L - \tilde{L}$ is precisely the excess degree vector of L. Choosing a decomposition $d^e = \sum \alpha_{ij} \mathbb{1}_{i,j}$, we write $L - \tilde{L} = \alpha_{ij} M_{ij}$, with M_{ij} constructed in the same way as the matrices L_e, but by choosing orthogonal matrices that sum to zero. The matrices M_e are also positive linear combinations of connection Laplacians, so $L = \sum_e L_e + \sum_{ij} M_{ij}$ is a positive linear combination of connection Laplacians.

□

3.4 KRON REDUCTION

Kron reduction is one of many names[2] given to the process of reducing an electrical network to an equivalent one on a given set of terminals. If a set U of nodes are left unconnected to the outside world (and hence obey Kirchoff's laws) we may view the circuit as a black box with only the remaining nodes $B = G \setminus U$ exposed. When we apply voltages x_B to these terminals, the remaining parameters will be determined by harmonic extension. As a result, the induced current flow at the terminals in B is equal to $L[B, B]x_B - L[B, U]x_U$, where x_U is the harmonic extension of x_B to U. But we can compute $x_U = L[U, U]^{-1} L[U, B]v_U$, so the linear operator

2 Other related terms include *response matrix*, *star-mesh transform*, *equivalent network*, and *transfer function*. Moving slightly further afield, the analogous concept for partial differential equations with boundary conditions is the *Dirichlet-to-Neumann operator*.

mapping voltages on B to current flows on B is $L[B,B] - L[B,U]L[U,U]^{-1}L[U,B]$. This is readily identified as the *Schur complement* $L/L[U,U]$ of $L[U,U]$ in L.

Surprisingly, the resulting matrix $L_B = L/L[U,U]$ is again a graph Laplacian. That is, the class of graph Laplacians is closed under Schur complements. The question then arises: can we perform Kron reduction on sheaves? Equivalently: is the class of sheaf Laplacians closed under Schur complements (with respect, of course, to some B for which harmonic extension is unique)?

This question is somewhat ill-posed so far. Which class of sheaf Laplacians are we talking about? What are the dimensions of stalks? As it turns out, the answer depends on which classes of sheaf Laplacians we choose.

Proposition 3.4.1. *If L is the degree 0 Laplacian of a sheaf with one-dimensional vertex stalks $\mathbb{R}$, then for any subset of vertices B with complement U, the Schur complement $L[B,B] - L[B,U]L[U,U]^{-1}L[U,B]$ is also a sheaf Laplacian whenever $L[U,U]$ is invertible.*

Proof. Recall the characterization of sheaves with one-dimensional vertex stalks from Section 3.3 as scaled diagonally dominant matrices. We first show that diagonally dominant matrices with positive diagonal are closed under Schur complement. These are equivalently the class of symmetric positive definite H-matrices: those whose *comparison matrices* $M(A)$, replacing diagonal elements with their absolute values and off-diagonal elements with their negative absolute values, can be written as $sI - B$ for some nonnegative matrix B. But it is an established fact that H-matrices and symmetric positive definite matrices are closed under Schur complements [JS05], so their intersection is as well.

Now, if a class $\mathcal{L}$ of matrices is closed under Schur complements, so is its diagonal scaling, consisting of matrices DLD for $L \in \mathcal{L}$ and D invertible and diagonal. This is for the simple reason that

$$\begin{aligned}
&(DLD)[B,B] - (DLD)[B,U](DLD)[U,U]^{-1}(DLD)[U,B] \\
&= D_B L[B,B] D_B - D_B L[B,U] D|_U D|_U^{-1} L[U,U]^{-1} D|_U^{-1} D_U L[U,B] D_B \\
&\qquad = D|_B \left(L[B,B] - L[B,U]L[U,U]^{-1}L[U,B] \right) D|_B
\end{aligned}$$

Thus, the class of scaled diagonally dominant matrices is closed under Schur complements, and hence Kron reduction can be applied to sheaves with all vertex stalks $\mathbb{R}$. □

However, sheaves whose vertex stalk dimensions vary can fail to have Kron reductions. For instance, the sheaf over $K_{1,3}$ in Figure 3.2 does not reduce to a sheaf on any graph on the three outer vertices. To see this, note that the space of

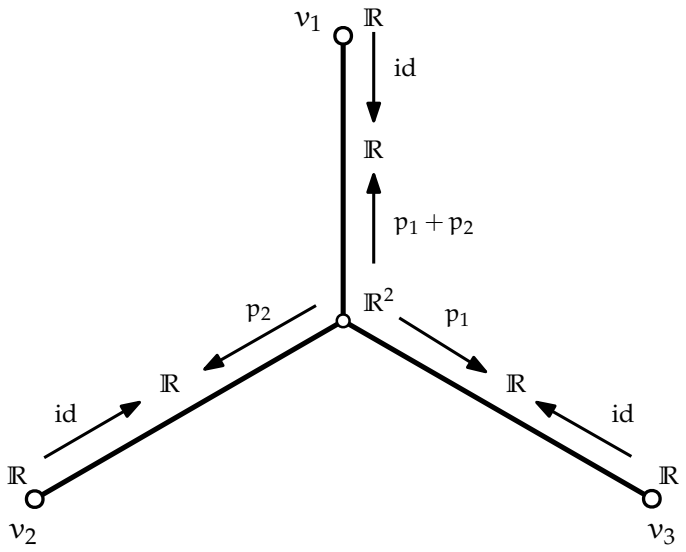


Figure 3.2: A sheaf which does not admit a Kron reduction.

global sections is two dimensional, and on the boundary vertices is spanned by the vectors

$$x_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Any graph on the three boundary vertices is a subgraph of K_3, and hence a sheaf on such a graph can be seen as a sheaf on K_3, so we consider any sheaf $\mathcal{G}$ on K_3 with vertex stalks $\mathbb{R}$. If x_1 is a section of $\mathcal{G}$, we must have $\mathcal{G}_{v_1 \trianglelefteq e_{13}} = 0$, where e_{13} is the edge between v_1 and v_3. Similarly, $\mathcal{G}_{v_2 \trianglelefteq e_{23}} = 0$. Further, if x_2 is also a section, we must have $\mathcal{G}_{v_1 \trianglelefteq e_{12}} = 0$ and $\mathcal{G}_{v_3 \trianglelefteq e_{23}} = 0$. But now because x_1 is a section, we must have $\mathcal{G}_{v_2 \trianglelefteq e_{12}} 1 = \mathcal{G}_{v_1 \trianglelefteq e_{12}} 1 = 0$, and similarly $\mathcal{G}_{v_3 \trianglelefteq e_{13}} 1 = \mathcal{G}_{v_1 \trianglelefteq e_{13}} 1 = 0$, so all restriction maps are zero, giving $\mathcal{G}$ a 3-dimensional space of global sections.

Even sheaves with constant-dimensional vertex stalks are not closed under Schur complements. Consider the Laplacian of a matrix-weighted graph, with a set of boundary vertices B. The question is whether the matrix $L_B = L[B, B] - L[B, S]L[S, S]^{-1}L[S, B]$ is the Laplacian of some sheaf. The kernel of L^B must include the space of constant cochains on B, since these extend with zero energy to the rest of the graph. Thus, if L_B is the Laplacian of a sheaf, its off-diagonal blocks must be symmetric, since this sheaf must have restriction maps of the form $\mathcal{F}_{v \trianglelefteq e} = \mathcal{F}_{u \trianglelefteq e}$. However, the off-diagonal blocks of L_B are $L[u, v] - L[u, S]L[S, S]^{-1}L[S, v]$, and while

$L[u,v]$ is symmetric, $L[u,S]L[S,S]^{-1}L[S,v]$ is not in general symmetric. (They are only guaranteed to be symmetric if $L[u,S] = L[v,S]$.)

On the other hand, the block norm diagonally dominant matrices of Section 3.3 are closed under Schur complements, as was shown in [Kyn+16]. As a result, one may perform Kron reduction on $O(n)$-bundles. The output is not necessarily a bundle, but will be representable as a cellular sheaf. This can be seen as a process yielding an $O(n)$-bundle on a multigraph with loops.

One way to interpret these results is that sheaves are more interesting than electrical circuits—they have interesting internal behavior that cannot always be predicted from their boundaries. Adding more nodes can increase the expressive power of a system described by a sheaf, even if those nodes are never observed.

4

SPECTRA OF SHEAF LAPLACIANS

Graph Laplacians are one of the principal points of interest in spectral graph theory, which seeks to understand the relationship between combinatorial properties of graphs and the eigenvalues of their associated matrices. Many results about the spectra of graph Laplacians have natural analogues for the spectra of Laplacians of sheaves on cell complexes. Often the difficulty in generalizing lies not in proving the generalized statement, but finding the correct generalization of graph-theoretic concepts to cellular sheaves.

4.1 PRELIMINARIES

An important tool in the spectral analysis of Laplacian matrices is the Courant-Fischer-Weyl theorem. This is a variational characterization of the eigenvalues and eigenvectors of a self-adjoint operator in terms of the *Rayleigh quotient*.

Definition 4.1.1 (Rayleigh quotient)**.** Let A be a self-adjoint operator on a Hilbert space $\mathbb{V}$, with $x \in \mathbb{V}$. The *Rayleigh quotient* of x with respect to A is

$$R_A(x) = \frac{\langle x, Ax\rangle}{\langle x, x\rangle}.$$

Theorem 4.1.1 (Courant-Fischer-Weyl)**.** *Let A be a self-adjoint operator on a finite-dimensional Hilbert space $\mathbb{V}$ of dimension n. If $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ are the eigenvalues of A, then*

$$\lambda_k = \max_{\substack{U \leqslant \mathbb{V} \\ \dim U = n-k+1}} \min_{x \in U} R_A(x) = \min_{\substack{U \leqslant \mathbb{V} \\ \dim U = k}} \max_{x \in U} R_A(x).$$

Proof. Because A is self-adjoint, it is unitarily equivalent to a diagonal matrix, so we may assume without loss of generality that A is diagonal, with diagonal entries $\lambda_1 \leqslant \cdots \leqslant \lambda_n$. Then

$$R_A(x) = \frac{\sum \lambda_i x_i^2}{\sum x_i^2}.$$

For a given $U \leqslant V$ of dimension $n - k + 1$, the intersection of U with the span of the first k basis vectors is nonempty, so $\min_{x \in U} R_A(x) \leqslant \lambda_k$. Similarly, for U of dimension k, $\max_{x \in U} R_A(x) \geqslant \lambda_k$. But it is straightforward to exhibit subspaces U for which the Rayleigh quotients attain these bounds exactly—simply take spans of basis vectors. This establishes the theorem. □

This minimax characterization of eigenvalues allows us to say things about the spectrum of a Laplacian matrix Δ in terms of its associated quadratic form $\langle x, \Delta x \rangle$.

A couple of basic results about the spectra of the operators Δ^k follow directly from discrete Hodge theory:

Proposition 4.1.1. *The nonzero elements of the spectrum of Δ^k are given by the disjoint union of the nonzero elements of the spectra of Δ^k_+ and Δ^k_-.*

Proof. The Hodge decomposition of $C^k(X; \mathcal{F})$ (Theorem 1.6.1) says that it is the orthogonal direct sum of $\ker \Delta^k$, $\operatorname{im} \Delta^k_-$, and $\operatorname{im} \Delta^k_+$. Since Δ^k_+ restricts to zero on $\operatorname{im} \Delta^k_-$ and Δ^k_- restricts to zero on $\operatorname{im} \Delta^k_+$, and both restrict to zero on $\ker \Delta^k$, we see that Δ^k is the orthogonal direct sum $0|_{\ker \Delta^k} \oplus \Delta^k_+|_{\operatorname{im} \Delta^k_+} \oplus \Delta^k_-|_{\operatorname{im} \Delta^k_-}$. As a result, the spectrum of Δ^k is the union of the spectra of these operators. □

Proposition 4.1.2. *The nonzero eigenvalues of Δ^k_+ and Δ^{k+1}_- are the same.*

Proof. Take the singular value decomposition $\delta^k = U \Sigma V^*$. Since $\Delta^k_+ = (\delta^k)^* \delta^k$ and $\Delta^{k+1}_- = \delta^k (\delta^k)^*$, we have unitary diagonalizations $\Delta^k_+ = V \Sigma^* \Sigma V^*$ and $\Delta^{k+1}_- = U \Sigma \Sigma^* U^*$. This implies that the eigenvalues of Δ^k_+ are the diagonal entries of $\Sigma^* \Sigma$, and the eigenvalues of Δ^{k+1}_- are the diagonal entries of $\Sigma \Sigma^*$. The nonzero entries of these matrices are the same. □

We can extract the eigenvectors of Δ^{k+1}_- from the eigenvectors of Δ^k_+. If v is an eigenvector of Δ^k_+, $\delta^k v$ is an eigenvector of Δ^{k-1}_-, since $\Delta^{k+1}_- \delta^k v = \delta^k \Delta^k_+ v = \lambda \delta^k v$.

Proposition 4.1.3. *Suppose $\mathcal{F}$ is a normalized sheaf on a simplicial complex. Then the eigenvalues of the degree k up-Laplacian Δ^k_+ of $\mathcal{F}$ are bounded above by $k + 2$.*

Proof. The Courant-Fischer-Weyl theorem implies that the largest eigenvalue of Δ^k_+ is equal to

$$\max_{x \in C^k(X; \mathcal{F})} \frac{\langle x, \Delta^k_+ x \rangle}{\langle x, x \rangle}$$

By the normalization condition on $\mathcal{F}$, we can write the denominator as $\langle x, x\rangle = \sum_{\dim \sigma = k} \langle \delta^k x_\sigma, \delta^k x_\sigma\rangle$, giving us the equivalent formula

$$\max_{x \perp \ker \delta^k} \frac{\langle \delta^k x, \delta^k x\rangle}{\sum_{\dim \sigma = k} \langle \delta^k x_\sigma, \delta^k x_\sigma\rangle} = \max_{x \perp \ker \delta^k} \frac{\sum_{\dim \tau = k+1} \sum_{\sigma, \sigma' \trianglelefteq_1 \tau} [\sigma : \tau][\sigma' : \tau]\langle \mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma, \mathcal{F}_{\sigma' \trianglelefteq \tau} x_{\sigma'}\rangle}{\sum_{\dim \sigma = k} \sum_{\sigma \trianglelefteq_1 \tau} \langle \mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma, \mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\rangle}$$

The Cauchy-Schwarz inequality implies that for $\sigma \neq \sigma'$,

$$\begin{aligned}[\sigma : \tau][\sigma' : \tau]\langle \mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma, \mathcal{F}_{\sigma' \trianglelefteq \tau} x_{\sigma'}\rangle &\leqslant \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\| \|\mathcal{F}_{\sigma' \trianglelefteq \tau} x_{\sigma'}\| \\ &\leqslant \frac{1}{2}\left(\|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2 + \|\mathcal{F}_{\sigma' \trianglelefteq \tau} x_{\sigma'}\|^2\right).\end{aligned}$$

If we substitute this in the inner sum of the numerator of the Rayleigh quotient, we get a bound of

$$\sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2 + \frac{1}{2} \sum_{\sigma \neq \sigma' \trianglelefteq_1 \tau} \left(\|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2 + \|\mathcal{F}_{\sigma' \trianglelefteq \tau} x_{\sigma'}\|^2\right).$$

Counting up all the terms of each type, we find that this is equal to

$$(k+2) \sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2.$$

Thus the Rayleigh quotient is bounded above by

$$\max_{x \perp \ker \delta^k} \frac{(k+2) \sum_{\dim \tau = k+1} \sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2}{\sum_{\dim \tau = k+1} \sum_{\sigma \trianglelefteq_1 \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} x_\sigma\|^2} = k+2.$$

□

This result is a generalization of a standard result about normalized graph Laplacians and a result of Horak and Jost about normalized Hodge Laplacians of simplicial complexes [HJ13]. The requirement that X be a simplicial complex is critical. This stipulation means that each k-cell has $k+1$ faces, which allows us to bound the sum in the numerator. In general, for a cell complex where each k-cell

has at most $b(k)$ faces, the spectrum of Δ_+^k for a normalized sheaf is bounded above by $b(k+1)$. A similar result about spectra of normalized factor width k matrices was given by Boman et al. [Bom+05], although their proof does not extend to block matrices.

4.2 EIGENVALUE INTERLACING

Definition 4.2.1. Let A, B be $n \times n$ matrices with real eigenvalues. Let $\lambda_1 \leqslant \lambda_2 \leqslant \cdots \leqslant \lambda_n$ be the eigenvalues of A and $\mu_1 \leqslant \mu_2 \leqslant \cdots \leqslant \mu_n$ be the eigenvalues of B. Extend these values to all integers by letting $\lambda_k = \lambda_1$ for $k < 1$ and $\lambda_k = \lambda_n$ for $k > n$.

We say the eigenvalues of A are (p, q)-*interlaced* with the eigenvalues of B if for all k, $\lambda_{k-p} \leqslant \mu_k \leqslant \lambda_{k+q}$.

Theorem 4.2.1. *Let A and B be positive semidefinite matrices, with $\operatorname{rank} B = t$. Then the eigenvalues of A are $(t, 0)$-interlaced with the eigenvalues of $A - B$.*

Proof. Let the eigenvalues of $A - B$ be $\mu_1 \leqslant \cdots \leqslant \mu_n$, and the eigenvalues of A be $\lambda_1 \leqslant \cdots \lambda_n$. We will use the Courant-Fischer-Weyl theorem to bound the eigenvalues of $A - B$. Note that the Rayleigh quotient $R_{A-B}(x)$ is equal to $R_A(x) - R_B(x)$. We then have

$$\begin{aligned}
\mu_k &= \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{x \in U} (R_A(x) - R_B(x)) \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{x \in U \cap \ker B} R_A(x) \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k-t}} \max_{x \in U} R_A(x) = \lambda_{k-t}.
\end{aligned}$$

The inequality in the third line comes by noting that $\dim U \cap \ker B$ has dimension at least $k - t$. The other side of the inequality is more straightforward:

$$\begin{aligned}
\lambda_k &= \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{x \in U} R_A(x) \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{x \in U} (R_A(x) - R_B(x)) = \mu_k.
\end{aligned}$$

□

The interlacing theorem gives results about the Laplacians of sheaves restricted to a subcomplex. Consider a $(k+1)$-dimensional complex X with a subcomplex

$A \subseteq X$ which contains all cells of dimension $\leqslant k$. We take the pullback $i^*\mathcal{F}$ over the inclusion map $i : A \to X$, and we compare the Laplacians $\Delta^k_{\mathcal{F}}$ and $\Delta^k_{i^*\mathcal{F}}$.

Our goal is to derive an interlacing relationship between the eigenvalues of $\Delta^k_{\mathcal{F}}$ and $\Delta^k_{i^*\mathcal{F}}$. To do this, we need to know the rank of $\Delta^k_{\mathcal{F}} - \Delta^k_{i^*\mathcal{F}}$. By interpreting this matrix as the Hodge Laplacian of another sheaf, we have a homological characterization. Let $\mathcal{G}$ be the sheaf on X with the same stalks as $\mathcal{F}$ but all restriction maps between cells in A set to zero. Then it is straightforward to check that $\Delta^k_{\mathcal{G}} = \Delta^k_{\mathcal{F}} - \Delta^k_{i^*\mathcal{F}}$, and applying Theorem 4.2.1 yields the following result:

Proposition 4.2.1. *The eigenvalues of $\Delta^k_{\mathcal{F}}$ are $(t, 0)$-interlaced with the eigenvalues of $\Delta^k_{i^*\mathcal{F}}$, where $t = \operatorname{codim} \mathcal{H}^k(X; \mathcal{G}) = \dim C^k(X; \mathcal{F}) - \dim H^k(X; \mathcal{G})$.*

Results about eigenvalue interlacing in spectral graph theory are typically stated in terms of the removal of an edge or vertex of a graph, but the formulation in terms of subcomplexes is more general, and allows a nice homological interpretation. There are also interlacing results about the normalized graph Laplacian. This is less straightforward to generalize to sheaves, since we have defined normalization as a property of the sheaf itself rather than its Laplacian. There is in general no simple relationship between the Laplacian of a cellular sheaf and the Laplacian of its normalization. However, for a sheaf on a graph, we did derive a formula similar to the one for the normalized graph Laplacian in Section 1.7.2. This gives us a result about eigenvalues of normalizations of the restriction of a sheaf onto a subcomplex.

Proposition 4.2.2. *Suppose $\mathcal{F}$ is a normalized sheaf on a graph X, with $A \subseteq X$ a subgraph. Let $\widetilde{i^*\mathcal{F}}$ be the normalization of the sheaf $i^*\mathcal{F}$. The eigenvalues of $L_{\mathcal{F}}$ are (t, t)-interlaced with the eigenvalues of $L_{\widetilde{i^*\mathcal{F}}}$, where $t = \operatorname{codim} \mathcal{H}^0(X; \mathcal{G})$.*

Proof. The Laplacian $L_{\widetilde{i^*\mathcal{F}}}$ written in terms of the same basis as $L_{\mathcal{F}}$ is equal to $D^{-1/2}_{i^*\mathcal{F}} L_{i^*\mathcal{F}} D^{-1/2}_{i^*\mathcal{F}}$, so that its Rayleigh quotient is

$$R_{L_{\widetilde{i^*\mathcal{F}}}}(x) = \frac{\langle x, D^{-1/2}_{i^*\mathcal{F}} L_{i^*\mathcal{F}} D^{-1/2}_{i^*\mathcal{F}} x\rangle}{\langle x, x\rangle} = \frac{\langle y, L_{i^*\mathcal{F}} y\rangle}{\langle y, D_{i^*\mathcal{F}} y\rangle} = \frac{\langle y, L_{\mathcal{F}} y\rangle - \langle y, L_{\mathcal{G}} y\rangle}{\langle y, y\rangle - \langle y, D_{\mathcal{G}} y\rangle}.$$

Letting λ_k be the kth largest eigenvalue of $L_{\mathcal{F}}$ and μ_k the kth largest eigenvalue of $L_{N(i^*\mathcal{F})}$, we have

$$\begin{aligned}
\mu_k &= \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{y \in U} \frac{\langle y, L_{\mathcal{F}} y\rangle - \langle y, L_{\mathcal{G}} y\rangle}{\langle y, y\rangle - \langle y, D_{\mathcal{G}} y\rangle} \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{y \in U \cap H^0(X;\mathcal{G})} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle - \langle y, D_{\mathcal{G}} y\rangle} \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k}} \max_{y \in U \cap H^0(X;\mathcal{G})} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle} \\
&\geqslant \min_{\substack{U \leqslant \mathbb{R}^n \\ \dim U = k-t}} \max_{y \in U} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle} = \lambda_{k-t}
\end{aligned}$$

and

$$\begin{aligned}
\mu_k &= \max_{\substack{U \leqslant \mathbb{R}^n \\ \dim U n-k+}} \min_{y \in U} \frac{\langle y, L_{\mathcal{F}} y\rangle - \langle y, L_{\mathcal{G}} y\rangle}{\langle y, y\rangle - \langle y, D_{\mathcal{G}} y\rangle} \\
&\leqslant \max_{\substack{U \leqslant \mathbb{R}^n \\ \dim U n-k+}} \min_{y \in U \cap H^0(X;\mathcal{G})} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle - \langle y, D_{\mathcal{G}} y\rangle} \\
&\leqslant \max_{\substack{U \leqslant \mathbb{R}^n \\ \dim U n-k+}} \min_{y \in U \cap H^0(X;\mathcal{G})} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle} \\
&\leqslant \max_{\substack{U \leqslant \mathbb{R}^n \\ \dim U n-k-t+}} \min_{y \in U} \frac{\langle y, L_{\mathcal{F}} y\rangle}{\langle y, y\rangle} = \lambda_{k+t}.
\end{aligned}$$

□

4.3 MORPHISMS

The existence of a sheaf morphism $\mathcal{F} \to \mathcal{G}$ does not in general give much information about the spectra of the corresponding sheaf Laplacians. However, unitary morphisms do offer some control over the relationships between sheaf Laplacians.

Proposition 4.3.1. *Let $\varphi : \mathcal{F} \to \mathcal{G}$ be a morphism of weighted sheaves on a regular cell complex X. If $\varphi^{k+1} : C^{k+1}(X;\mathcal{F}) \to C^{k+1}(X;\mathcal{G})$ is unitary, then $(\Delta_+^k)_{\mathcal{F}} = (\varphi^k)^*(\Delta_+^k)_{\mathcal{G}}\varphi^k$.*

Proof. Since sheaf morphisms commute with the coboundary maps, we have

$$(\delta_{\mathcal{F}}^k)^*(\varphi^{k+1})^*\varphi^{k+1}\delta_{\mathcal{F}}^k = (\varphi^k)^*(\delta_{\mathcal{G}}^k)^*\delta_{\mathcal{G}}^k\varphi^k = (\varphi^k)^*(\Delta_+^k)_{\mathcal{G}}\varphi^k.$$

As a result, if $(\varphi^{k+1})^*\varphi^{k+1} = \mathrm{id}$, we have $(\Delta^k_+)_{\mathcal{F}} = (\varphi^k)^*(\Delta^k_+)_{\mathcal{G}}\varphi^k$, and this condition holds if φ^{k+1} is unitary. □

Analogously, we also have

Proposition 4.3.2. *Let $\varphi : \mathcal{F} \to \mathcal{G}$ be a morphism of weighted sheaves on a regular cell complex X. If $\varphi^{k-1} : C^{k-1}(X;\mathcal{F}) \to C^{k-1}(X;\mathcal{G})$ is unitary, then $(\Delta^k_-)_{\mathcal{F}} = (\varphi^k)^*(\Delta^k_-)_{\mathcal{G}}\varphi^k$.*

Thus, if both φ^{k+1} and φ^{k-1} are unitary, we have $\Delta^k_{\mathcal{F}} = (\varphi^k)^*\Delta^k_{\mathcal{G}})\varphi^k$. Even if φ^k is not unitary, we have a bound

$$\lambda_{\max}(\Delta^k_{\mathcal{F}}) \leqslant \lambda_{\max}(\Delta^k_{\mathcal{G}})\|\varphi^k\|^2,$$

coming from the ℓ^2 operator norm. Of course, if φ^k is unitary, the spectra of the two operators are identical.

Naturally, the direct sum of sheaves is spectrally well-behaved.

Proposition 4.3.3. *Let $\mathcal{F}$ and $\mathcal{G}$ be sheaves on X. The Laplacians $\Delta^k_{\mathcal{F}\oplus\mathcal{G}}$ naturally decompose into an orthogonal direct sum $\Delta^k_{\mathcal{F}} \oplus \Delta^k_{\mathcal{G}}$.*

Proof. The complex of cochains $C^\bullet(X;\mathcal{F}\oplus\mathcal{G})$ naturally decomposes into an orthogonal direct sum $C^\bullet(X;\mathcal{F}) \oplus C^\bullet(X;\mathcal{G})$. Since the coboundary decomposes with respect to this direct sum, so do all the Laplacians. □

In particular, this means that the spectrum of $\Delta^k_{\mathcal{F}\oplus\mathcal{G}}$ is the disjoint union of the spectra of $\Delta^k_{\mathcal{F}}$ and $\Delta^k_{\mathcal{G}}$.

The tensor product of sheaves is much harder to get a spectral handle on. Difficulties arise even in predicting the dimension of the space of global sections. If $\mathcal{F}$ is an $O(1)$-bundle, then there exists a sheaf $\mathcal{F}^{-1}$ such that $\mathcal{F}\otimes\mathcal{F}^{-1} \simeq \underline{\mathbf{k}}$. This means that if $\mathcal{F}$ has no global sections, there are some sheaves $\mathcal{G}$ such that $\mathcal{F}\otimes\mathcal{G}$ has global sections, while there are others for which it does not. Whether this happens seems to be very difficult to predict purely from the Laplacian spectrum.

Another point of difficulty here is that the cochain complex of $\mathcal{F}\otimes\mathcal{G}$ does not decompose nicely into anything naturally acted on by the component coboundary operators. That is, $C^k(X;\mathcal{F}\otimes\mathcal{G})$ is a direct sum of tensor products of stalks, and direct sums do not distribute over tensor products.

More interesting results are possible when we investigate the effects of morphisms of cell complexes and their associated sheaf operations.

Proposition 4.3.4. *Let $f : X \to Y$ be a locally injective cellular map. If $\mathcal{F}$ is a sheaf on X, the degree-k up-Laplacian $(\Delta^k_+)_{f_*\mathcal{F}}$ corresponding to $f_*\mathcal{F}$ on Y is unitarily equivalent to the degree k coboundary Laplacian $(\Delta^k_+)_{\mathcal{F}}$ of $\mathcal{F}$.*

Proof. The morphism f induces a set of isometries $f_k : C^k(X;\mathcal{F}) \to C^k(Y;f_*\mathcal{F})$, which are given on stalks by the inclusion $f_\sigma : \mathcal{F}(\sigma) \to f_*\mathcal{F}(f(\sigma)) = \bigoplus_{f(\tau)=f(\sigma)} \mathcal{F}(\tau)$. These stalkwise maps commute with the restriction maps of $\mathcal{F}$ and $f_*\mathcal{F}$, so f_k commutes with the coboundary maps. Therefore,

$$(\Delta^k_+)_{f_*\mathcal{F}} = (\delta^k_{f_*\mathcal{F}})^*\delta^k_{f_*\mathcal{F}} = (\delta^k_{f_*\mathcal{F}})^* f^*_{k+1} f_{k+1}\delta^k_{f_*\mathcal{F}} = f^*_k(\delta^k_{\mathcal{F}})^*\delta^k_{\mathcal{F}} f_k = f^*_k(\Delta^k_+)_{\mathcal{F}}.$$

□

Corollary. *The sheaves $\mathcal{F}$ and $f_*\mathcal{F}$ are isospectral for the up-Laplacian.*

Proposition 4.3.5. *Let $p : E \to X$ be a covering map of cell complexes, with $\mathcal{F}$ a sheaf on X. Then for any k, the spectrum of any Laplacian of $\mathcal{F}$ is contained in the spectrum of the corresponding Laplacian of $p^*\mathcal{F}$.*

Proof. There is a lifting map $p^* : C^\bullet(X;\mathcal{F}) \to C^\bullet(E;p^*\mathcal{F})$ which takes $(x_\sigma)_{\sigma\in X} \in C^k(X;\mathcal{F})$ to $(x_{p(\tau)})_{\tau\in E} \in C^k(E;p^*\mathcal{F})$. We will denote its adjoint $(p^*)^*$ by simply p; this map is given by $(p(x))_\sigma = \sum_{\sigma'\in p^{-1}(\sigma)} x_{\sigma'}$. The map p^* commutes with both δ and δ^*. It commutes with δ because it commutes with every restriction map (just like the isometries in the proof of Proposition 4.3.4). The commutativity with δ^* relies on the fact that p is a covering map and the behavior of the pullback sheaf.

For $y \in C^k(E;p^*\mathcal{F})$ and $x \in C^{k+1}(X;\mathcal{F})$, we have

$$\begin{aligned}
\langle y, \delta^* p^* x\rangle = \langle \delta y, p^* x\rangle &= \sum_{\substack{\sigma',\tau'\in E\\ \sigma' \trianglelefteq_1 \tau'}} [\sigma' : \tau']\langle p^*\mathcal{F}_{\sigma' \trianglelefteq \tau'}(y_{\sigma'}), (p^*x)_{\tau'}\rangle \\
&= \sum_{\substack{\sigma,\tau\in X\\ \sigma \trianglelefteq_1 \tau}} [\sigma : \tau] \sum_{\sigma'\in f^{-1}(\sigma)} \langle \mathcal{F}_{\sigma\trianglelefteq\tau}(y_{\sigma'}), x_\tau\rangle \\
&= \sum_{\substack{\sigma,\tau\in X\\ \sigma \trianglelefteq_1 \tau}} [\sigma : \tau]\langle \mathcal{F}_{\sigma\trianglelefteq\tau}(py)_\sigma, x_\tau\rangle \\
&= \langle \delta p y, x\rangle = \langle y, p^*\delta^* x\rangle
\end{aligned}$$

Suppose now that $(\Delta^k_+)_{\mathcal{F}} x = \lambda x$. Then $(\Delta^k_+)_{p^*\mathcal{F}} p^* x = (\delta^k_{p^*\mathcal{F}})^*\delta^k_{p^*\mathcal{F}} p^* x = p^*(\delta^k_{\mathcal{F}})^*\delta^k_{\mathcal{F}} x = p^*(\Delta^k_+)_{\mathcal{F}} = \lambda p^* x$, so λ is an eigenvalue of $(\Delta^k_+)_{p^*\mathcal{F}}$ with eigenvector p^*x. A similar argument holds for Δ^k_- and Δ^k. □

Pullbacks over other sorts of maps can also yield some control over the spectrum. For instance, if we have a dimension-preserving map with uniform fiber sizes, we can bound the smallest nontrivial eigenvalue:

Proposition 4.3.6. *Suppose $f : Y \to X$ is a cellular map with $\dim f(\sigma) = \dim \sigma$ for all cells $\sigma \in Y$, and further, that for each cell σ of X of dimension d, $\left|f^{-1}(\sigma)\right| = \ell_d$ is constant.*

If $\mathcal{F}$ is a weighted sheaf on X, with $\lambda(\mathcal{F})$ the smallest nontrivial eigenvalue of $(\Delta_+^k)_{\mathcal{F}}$, then $\lambda(\mathcal{F}) \geqslant \frac{\ell_k}{\ell_{k+1}}(f^*\mathcal{F})$.

Proof. Let x be an eigenvector of $(\Delta_+^k)_{\mathcal{F}}$ corresponding to $\lambda(\mathcal{F})$. Since every fiber of f is of the same size, the lifting map $f_k^* : C^k(X;\mathcal{F}) \to C^k(Y; f^*\mathcal{F})$ preserves the inner product up to a scaling. That is, $\langle f^*y, f^*z\rangle = \ell_k\langle y, z\rangle$ for $y, z \in C^k(X;\mathcal{F})$. As a result, f^*x is orthogonal to f^*y for any $y \in \ker(\Delta_+^k)_{\mathcal{F}}$. Therefore, we have

$$\lambda(\mathcal{F}) = \frac{\langle \delta_{\mathcal{F}} x, \delta_{\mathcal{F}} x\rangle}{\langle x, x\rangle} = \frac{\ell_k\langle f^*\delta_{\mathcal{F}} x, f^*\delta_{\mathcal{F}} x\rangle}{\ell_{k+1}\langle f^*x, f^*x\rangle} = \frac{\ell_k\langle \delta_{f^*\mathcal{F}} f^*x, \delta_{f^*\mathcal{F}} f^*x\rangle}{\ell_{k+1}\langle f^*x, f^*x\rangle} \geqslant \frac{\ell_k}{\ell_{k+1}}\lambda(f^*\mathcal{F}).$$

□

A covering map of complexes $p : E \to X$ may be equivalently thought of as a fiber bundle over X with discrete stalks of size n, or a locally constant sheaf with values in the category of sets. We can compose this locally constant sheaf with the free vector space functor to get a sheaf on X; this sheaf is equivalent to the sheaf $p_*\underline{\mathbf{k}}$. To understand better what happens to the spectrum of $p^*\mathcal{F}$ for $\mathcal{F}$ a sheaf on X, note that by Proposition 4.3.4, this spectrum is the same as that of $p_*p^*\mathcal{F}$. It is straightforward to verify that $p_*p^*\mathcal{F} \simeq \mathcal{F} \otimes p_*\underline{\mathbf{k}}$. Further, $p_*\underline{\mathbf{k}}$ decomposes according to the representation of the structure group of the bundle on $\mathbf{k}^n$. That is, if the representation has a direct sum decomposition, so does $p_*\underline{\mathbf{k}}$. For instance, when $n = 2$, the structure group is $\mathbb{Z}/2$ and the representation is the canonical representation on $\mathbf{k}^2$, which decomposes into the direct sum of the trivial representation and the sign representation. Thus $p_*\underline{\mathbf{k}} \simeq \underline{\mathbf{k}} \oplus \tilde{\underline{\mathbf{k}}}$, where $\tilde{\underline{\mathbf{k}}}$ is an $O(1)$-bundle. The sign of each 1-cell corresponds to whether the fibers "switch places" over that cell. Thus, the spectrum of $p^*\mathcal{F}$ is the union of the spectra of $\mathcal{F}$ and $\mathcal{F} \otimes \tilde{\underline{\mathbf{k}}}$. For the case of a constant sheaf on a graph, we see that the "new eigenvalues" of the cover are the eigenvalues of a signed graph, where the signing corresponds to the cover. This fact, with less sheaf-theoretic cover, has been used to great effect in understanding the spectrum of a double cover of a graph [BLo6].

4.4 PRODUCTS

Definition 4.4.1. If $\mathcal{F}$ and $\mathcal{G}$ are sheaves on X and Y, respectively, their *outer product* is the sheaf $\mathcal{F} \boxtimes \mathcal{G} = \pi_X^*\mathcal{F} \otimes \pi_Y^*\mathcal{G}$ on $X \times Y$.

Unraveling this definition reveals that $(\mathcal{F} \boxtimes \mathcal{G})(\sigma \times \tau) = \mathcal{F}(\sigma) \otimes \mathcal{F}(\tau)$ and $(\mathcal{F} \boxtimes \mathcal{G})_{\sigma\times\tau \trianglelefteq \sigma'\times\tau'} = \mathcal{F}_{\sigma \trianglelefteq \sigma'} \otimes \mathcal{G}_{\tau \trianglelefteq \tau'}$. If $\mathcal{F}$ and $\mathcal{G}$ are weighted, there is a clear canonical weighting on $\mathcal{F} \boxtimes \mathcal{G}$. This is the analogue for sheaves on graphs of the Cartesian product of graphs.

Proposition 4.4.1. *If $L_{\mathcal{F}}$ and $L_{\mathcal{G}}$ are the degree-0 Laplacians of $\mathcal{F}$ and $\mathcal{G}$, then the degree-0 Laplacian of $\mathcal{F} \boxtimes \mathcal{G}$ is $L_{\mathcal{F}\boxtimes\mathcal{G}} = \mathrm{id}_{C^0(X;\mathcal{F})} \otimes L_{\mathcal{G}} + L_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})}$.*

Proof. The vector space $C^1(X \times Y; \mathcal{F} \boxtimes \mathcal{G})$ has an orthogonal decomposition into two subspaces: one generated by stalks $\mathcal{F}(v) \otimes \mathcal{G}(e)$ with v a vertex of X and e an edge of Y, and the other generated by stalks of the dual form $\mathcal{F}(e) \otimes \mathcal{G}(v)$. This is an isomorphism

$$C^1(X \times Y; \mathcal{F} \boxtimes \mathcal{G}) \simeq (C^0(X;\mathcal{F}) \otimes C^1(Y;\mathcal{G})) \oplus (C^1(X;\mathcal{F}) \otimes C^0(Y;\mathcal{G})).$$

The coboundary map of $\mathcal{F} \boxtimes \mathcal{G}$ is, with respect to this decomposition on C^1, the block matrix

$$\delta_{\mathcal{F}\boxtimes\mathcal{G}} = \begin{bmatrix} \mathrm{id}_{C^0(X;\mathcal{F})} \otimes \delta_{\mathcal{G}} \\ \delta_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})} \end{bmatrix}.$$

It is then straightforward to compute

$$\begin{aligned} L_{\mathcal{F}\boxtimes\mathcal{G}} = \delta^*_{\mathcal{F}\boxtimes\mathcal{G}}\delta_{\mathcal{F}\boxtimes\mathcal{G}} &= \mathrm{id}_{C^0(X;\mathcal{F})} \otimes \delta^*_{\mathcal{G}}\delta_{\mathcal{G}} + \delta^*_{\mathcal{F}}\delta_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})} \\ &= \mathrm{id}_{C^0(X;\mathcal{F})} \otimes L_{\mathcal{G}} + L_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})}. \end{aligned}$$

□

Corollary. *The eigenvalues of $L_{\mathcal{F}\boxtimes\mathcal{G}}$ are sums $\lambda + \mu$, where λ is an eigenvalue of $L_{\mathcal{F}}$ and μ is an eigenvalue of $L_{\mathcal{G}}$.*

Proof. If $L_{\mathcal{F}}v_{\mathcal{F}} = \lambda v_{\mathcal{F}}$ and $L_{\mathcal{G}}v_{\mathcal{G}} = \mu v_{\mathcal{G}}$, then $L_{\mathcal{F}\boxtimes\mathcal{G}}v_{\mathcal{F}} \otimes v_{\mathcal{G}} = v_{\mathcal{F}} \otimes \mu v_{\mathcal{G}} + \lambda v_{\mathcal{F}} \otimes v_{\mathcal{G}} = (\lambda + \mu)v_{\mathcal{F}} \otimes v_{\mathcal{G}}$. Thus the spectrum of $L_{\mathcal{F}\boxtimes\mathcal{G}}$ contains sums of the form $\lambda + \mu$. But this operator acts on the vector space $C^0(X;\mathcal{F}) \otimes C^0(Y;\mathcal{G})$, which has dimension $\dim(C^0(X;\mathcal{F}))\dim(C^0(Y;\mathcal{G}))$, and we have exhibited this many eigenvectors with the corresponding eigenvalues, so we have described the entire spectrum of $L_{\mathcal{F}\boxtimes\mathcal{G}}$. □

The higher degree coboundary maps of $\mathcal{F} \boxtimes \mathcal{G}$ are more complicated. For instance, in degree 1 we have

$$\delta^1_{\mathcal{F}\boxtimes\mathcal{G}} = \begin{bmatrix} \mathrm{id}_{C^0(X;\mathcal{F})} \otimes \delta^1_{\mathcal{G}} & 0 \\ \delta^0_{\mathcal{F}} \otimes \mathrm{id}_{C^1(Y;\mathcal{G})} & \mathrm{id}_{C^1(X;\mathcal{G})} \otimes \delta^0_{\mathcal{G}} \\ 0 & \delta^1_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})} \end{bmatrix}.$$

This then implies that the degree-1 up-Laplacian is equal to

$$\begin{bmatrix} \mathrm{id}_{C^0(X;\mathcal{F})} \otimes (\Delta^1_+)_{\mathcal{G}} + (\Delta^0_+)_{\mathcal{F}} \otimes \mathrm{id}_{C^1(Y;\mathcal{G})} & (\delta^0_{\mathcal{F}})^* \otimes \delta^0_{\mathcal{G}} \\ \delta^0_{\mathcal{F}} \otimes (\delta^0_{\mathcal{G}})^* & \mathrm{id}_{C^1(X;\mathcal{F})} \otimes (\Delta^0_+)_{\mathcal{G}} + (\Delta^1_+)_{\mathcal{F}} \otimes \mathrm{id}_{C^0(Y;\mathcal{G})} \end{bmatrix}.$$

The extra terms make the spectrum harder to compute. When X is a graph, $(\Delta^1_+)_{\mathcal{F}}$ and $(\Delta^1_+)_{\mathcal{G}}$ are zero, and we can compute the spectrum of $(\Delta^1_+)_{\mathcal{F}\boxtimes\mathcal{G}}$. (Note that this is not the zero matrix, since $X \times Y$ has 2-dimensional cells which are products of 1-dimensional cells.)

Proposition 4.4.2. *Let X and Y be graphs, with $X \times Y$ a 2-dimensional cell complex, and $\mathcal{F}$ and $\mathcal{G}$ sheaves on X and Y. If $v_{\mathcal{F}}$ is an eigenvector of $(\Delta^0_+)_{\mathcal{F}}$ with eigenvalue λ and $v_{\mathcal{G}}$ an eigenvector of $(\Delta^0_+)_{\mathcal{G}}$ with eigenvalue μ, then the vector*

$$v_{\mathcal{F}\boxtimes\mathcal{G}} = \begin{bmatrix} \sqrt{\frac{\lambda}{\mu}} v_{\mathcal{F}} \otimes \delta^0_{\mathcal{G}} v_{\mathcal{G}} \\ \sqrt{\frac{\mu}{\lambda}} \delta^0_{\mathcal{F}} v_{\mathcal{F}} \otimes v_{\mathcal{G}} \end{bmatrix}$$

is an eigenvector of $(\Delta^1_+)_{\mathcal{F}\boxtimes\mathcal{G}}$ with eigenvalue $\lambda + \mu$.

Proof. We compute

$$\begin{aligned} (\Delta^1_+)_{\mathcal{F}\boxtimes\mathcal{G}} v_{\mathcal{F}\boxtimes\mathcal{G}} &= \begin{bmatrix} (\Delta^0_+)_{\mathcal{F}} \otimes \mathrm{id}_{C^1(Y;\mathcal{G})} & (\delta^0_{\mathcal{F}})^* \otimes \delta^0_{\mathcal{G}} \\ \delta^0_{\mathcal{F}} \otimes (\delta^0_{\mathcal{G}})^* & \mathrm{id}_{C^1(X;\mathcal{F})} \otimes (\Delta^0_+)_{\mathcal{G}} \end{bmatrix} \begin{bmatrix} \sqrt{\frac{\lambda}{\mu}} v_{\mathcal{F}} \otimes \delta^0_{\mathcal{G}} v_{\mathcal{G}} \\ \sqrt{\frac{\mu}{\lambda}} \delta^0_{\mathcal{F}} v_{\mathcal{F}} \otimes v_{\mathcal{G}} \end{bmatrix} \\ &= \begin{bmatrix} \left(\sqrt{\frac{\lambda}{\mu}} + \sqrt{\frac{\mu}{\lambda}}\right) \lambda v_{\mathcal{F}} \otimes \delta^0_{\mathcal{G}} v_{\mathcal{G}} \\ \left(\sqrt{\frac{\lambda}{\mu}} + \sqrt{\frac{\mu}{\lambda}}\right) \mu \delta^0_{\mathcal{F}} v_{\mathcal{F}} \otimes v_{\mathcal{G}} \end{bmatrix} = (\lambda + \mu) \begin{bmatrix} \sqrt{\frac{\lambda}{\mu}} v_{\mathcal{F}} \otimes \delta^0_{\mathcal{G}} v_{\mathcal{G}} \\ \sqrt{\frac{\mu}{\lambda}} \delta^0_{\mathcal{F}} v_{\mathcal{F}} \otimes v_{\mathcal{G}} \end{bmatrix}. \end{aligned}$$

□

We can also get a similar result by recalling that the nonzero spectrum of $(\delta^1_{\mathcal{F}\boxtimes\mathcal{G}})^* \delta^1_{\mathcal{F}\boxtimes\mathcal{G}}$ is equal to the nonzero spectrum of $\delta^1_{\mathcal{F}\boxtimes\mathcal{G}} (\delta^1_{\mathcal{F}\boxtimes\mathcal{G}})^*$. This second term is easy to compute for X and Y graphs:

$$(\Delta^2_-)_{\mathcal{F}\boxtimes\mathcal{G}} = (\Delta^1_-)_{\mathcal{F}} \otimes \mathrm{id}_{C^1(Y;\mathcal{G})} + \mathrm{id}_{C^1(X;\mathcal{G})} \otimes (\Delta^1_-)_{\mathcal{G}}.$$

But then we know that the nonzero eigenvalues of $(\Delta^1_-)_{\mathcal{F}}$ are the same as those of $(\Delta^0_+)_{\mathcal{F}}$, we get a correspondence between nonzero eigenvalues of $(\Delta^1_+)_{\mathcal{F}\boxtimes\mathcal{G}}$ and sums of nonzero eigenvalues of $(\Delta^0_+)_{\mathcal{F}}$ and $(\Delta^0_+)_{\mathcal{G}}$.

As we move to higher dimensions, we can decompose

$$C^k(X \times Y; \mathcal{F} \boxtimes \mathcal{G}) = \bigoplus_{i=0}^{k} C^i(X; \mathcal{F}) \otimes C^{k-i}(Y; \mathcal{G})$$

and the matrix of $\delta^k_{\mathcal{F}\boxtimes\mathcal{G}}$ has progressively more blocks with more complex formulas, preventing straightforward calculation of the Laplacian eigenvalues.

5

SHEAF APPROXIMATIONS

If a cellular sheaf models some concrete physical system, it may be useful in certain situations to be able to approximate this sheaf in some sense. This is particularly interesting when we are designing a system whose complexity depends on the complexity of an underlying sheaf. Constructing a simple approximation would then mean a more efficient design. Here we will explore one notion of approximation for cellular sheaves.

5.1 SHEAF APPROXIMATIONS

Definition 5.1.1. Let $\mathcal{F}$ be a sheaf on a regular cell complex X. A k-*approximation* of $\mathcal{F}$ is a sheaf $\mathcal{G}$ on X together with a morphism $a : \mathcal{F} \to \mathcal{G}$ such that

1. a is an isomorphism on stalks over cells of dimension $\leqslant k$, and a surjection on stalks over cells of dimension $k+1$
2. $H^i(a) : H^i(X;\mathcal{F}) \to H^i(X;\mathcal{G})$ is an isomorphism for $i \leqslant k$.

One motivation for this definition is its clear analogy to approximation of spaces by weakly homotopy equivalent CW complexes. One way to prove that every space Y is weakly homotopy equivalent to a CW complex is to construct a sequence of skeleta X^k together with maps $X^k \to Y$ which induce isomorphisms $\pi_i(X^k) \to \pi_i(Y)$ for $i < k$ and are surjective for $i = k$. Using cohomology means that the approximating maps will go in the other direction, and since we are approximating not the space but an algebraic structure associated thereto, we can control its behavior at individual points as well.

Another motivation is more concrete in nature. A cellular sheaf of vector spaces represents a system of linear constraints for data on a space. The cohomology of a sheaf represents classes of solutions to certain equations parameterized by the space. If we are only interested in these constraints up to a certain dimension of cell, we may be able to ignore significant amounts of information in our original sheaf, simplifying computation and communication. The simplest example of this is a 0-approximation of a sheaf $\mathcal{F}$ on a graph. A 0-approximation to $\mathcal{F}$ has the same space of global sections as $\mathcal{F}$, but may have significantly lower-dimensional stalks over edges. If the graph represents a distributed system, this reduction in dimension corresponds to a reduction in connection complexity.

Definition 5.1.1 ignores cells of dimension greater than $k+1$, so when considering k-approximations we will assume without loss of generality that X is $(k+1)$-dimensional. Most of the difficulties in understanding sheaf approximations are evident in the case $k = 0$, so we will focus on that case. For each 1-cell e with faces u and v, a 0-approximation $a : \mathcal{F} \to \mathcal{G}$ defines a commuting diagram

$$\begin{array}{ccc}
\mathcal{F}(v) & \xrightarrow{\simeq} & \mathcal{G}(v) \\
{\scriptstyle \mathcal{F}_{v \trianglelefteq e}}\downarrow & & \downarrow{\scriptstyle \mathcal{G}_{v \trianglelefteq e}} \\
\mathcal{F}(e) & \xrightarrow{a_e} & \mathcal{G}(e) \\
{\scriptstyle \mathcal{F}_{u \trianglelefteq e}}\uparrow & & \uparrow{\scriptstyle \mathcal{G}_{u \trianglelefteq e}} \\
\mathcal{F}(u) & \xrightarrow{\simeq} & \mathcal{G}(u)
\end{array}$$

The approximation $\mathcal{G}$ is determined up to isomorphism by the choice of $\mathcal{G}(e)$ and a_e for every edge. Since a_e is surjective, $\mathcal{G}(e) \simeq \mathcal{F}(e)/\ker a_e$, so the structure of $\mathcal{G}$ is determined by the choice of $\ker a_e$ for every e. Thus, we may think of a 0-approximation as being given by a choice of a subspace of $\mathcal{F}(e)$ to discard. A similar characterization holds in higher dimensions: a k-approximation is given by a choice of a subspace of $\mathcal{F}(\sigma)$ to discard for every $(k+1)$-cell σ.

The following result gives us some information about when an approximation on a subcomplex extends to the entire complex, and can be used to glue together approximations.

Proposition 5.1.1. *Let $\mathcal{F}$ be a sheaf on a $(k+1)$-dimensional complex X, and let $Y \subseteq X$ be a subcomplex such that the induced map $H^k(X;\mathcal{F}) \to H^k(Y;\mathcal{F})$ is an isomorphism. If $a : \mathcal{F} \to \mathcal{G}$ is a morphism of sheaves on X which is an isomorphism on stalks over cells of dimension $\leqslant k$, such that $a|_Y : \mathcal{F}|_Y \to \mathcal{G}|_Y$ is a k-approximation and the map $H^k(X,Y;\mathcal{F}) \to H^k(X,Y;\mathcal{G})$ induced by the sheaf morphism a is an isomorphism, then $a : \mathcal{F} \to \mathcal{G}$ is a k-approximation.*

Proof. To show that $a : \mathcal{F} \to \mathcal{G}$ is a k-approximation, we only need to check that $H^k a : H^k(X;\mathcal{F}) \to H^k(X;\mathcal{G})$ is an isomorphism. Consider the two long exact sequences for relative sheaf cohomology of $\mathcal{F}$ and $\mathcal{G}$. We have a commutative diagram

$$\begin{array}{ccccccc}
H^k(X,Y;\mathcal{F}) & \xrightarrow{0} & H^k(X;\mathcal{F}) & \xrightarrow{\simeq} & H^k(Y;\mathcal{F}) & \longrightarrow & 0 \\
\downarrow{\scriptstyle \simeq} & & \downarrow{\scriptstyle H^k a} & & \downarrow{\scriptstyle \simeq} & & \downarrow \\
H^k(X,Y;\mathcal{G}) & \longrightarrow & H^k(X;G) & \longrightarrow & H^k(Y;G) & \longrightarrow & \operatorname{coker}
\end{array}$$

where each row is exact. Because a is an isomorphism on k-dimensional stalks, $H^k a$ must be injective, so we only need to show that $H^k a$ is surjective. But this follows from applying the five lemma to this commutative diagram. □

5.2 APPROXIMATIONS TO THE CONSTANT SHEAF

The first sheaf we will consider approximating is the constant sheaf. Consider the constant sheaf $\underline{\mathbb{V}}$ with stalk $\mathbb{V}$ on a graph G. If $\mathcal{F}$ is a 0-approximation to $\underline{\mathbb{V}}$, it is isomorphic to a sheaf with vertex stalks $\mathbb{V}$ such that for every edge e between vertices u and v, $\mathcal{F}_{u \trianglelefteq e} = \mathcal{F}_{v \trianglelefteq e}$. This is straightforward, since a_v defines an isomorphism between $\mathbb{V}$ and $\mathcal{F}(v)$ for all v. For an edge e, we have the diagram

$$\begin{array}{ccc}
\mathbb{V} & \xrightarrow{\mathrm{id}} & \mathbb{V} \\
{\scriptstyle \mathrm{id}}\downarrow & & \downarrow{\scriptstyle \mathcal{F}_{v \trianglelefteq e}} \\
\mathbb{V} & \xrightarrow{a_e} & \mathcal{F}(e) \\
{\scriptstyle \mathrm{id}}\uparrow & & \uparrow{\scriptstyle \mathcal{F}_{u \trianglelefteq e}} \\
\mathbb{V} & \xrightarrow{\mathrm{id}} & \mathbb{V}
\end{array}$$

which commutes if and only if $\mathcal{F}_{u \trianglelefteq e} = \mathcal{F}_{v \trianglelefteq e}$. This shows that degree-0 approximations to the constant sheaf are matrix-weighted graphs, with $W_{uv} = \mathcal{F}^*_{u \trianglelefteq e} \mathcal{F}_{v \trianglelefteq e}$. However, they satisfy an additional constraint: they have no nonconstant global sections. This constraint is nontrivial.

By the considerations for general approximations, $\mathcal{F}(e) \simeq \mathbb{V}/\ker a_e$. This suggests that we might construct an approximation to $\underline{\mathbb{V}}$ by choosing some $K_e \leqslant \mathbb{V}$ to be the kernel of a_e for each edge. If these kernels are chosen correctly, $H^0(G; \underline{\mathbb{V}})$ will be preserved. Choosing a collection $\{K_e\}$ that produces an approximation to $\underline{\mathbb{V}}$ is a subtle problem, but there are some obvious constraints.

Proposition 5.2.1. *If a collection $\{K_e\}$ determine an approximation $\mathcal{F}$ to the constant sheaf on a connected graph G, then for any set C of edges whose removal disconnects G, $\bigcap_{e \in C} K_e = 0$.*

Proof. Suppose there exists some nonzero vector $x \in \bigcap_{e \in C} K_e$. Removing the set of edges C separates G into at least two components. Pick a component G_1, and define the 0-cochain y by $y_v = x$ for $v \in G_1$ and $y_v = 0$ otherwise. This cochain is in $H^0(G; \mathcal{F})$: for every edge e not in C, $(\delta y)_e = 0$ because y is constant across that edge, and for every edge $e \in C$, $\mathcal{F}_{v \trianglelefteq e} y_v = 0$ because either $y_v = x \in K_e$ or $y_v = 0$. □

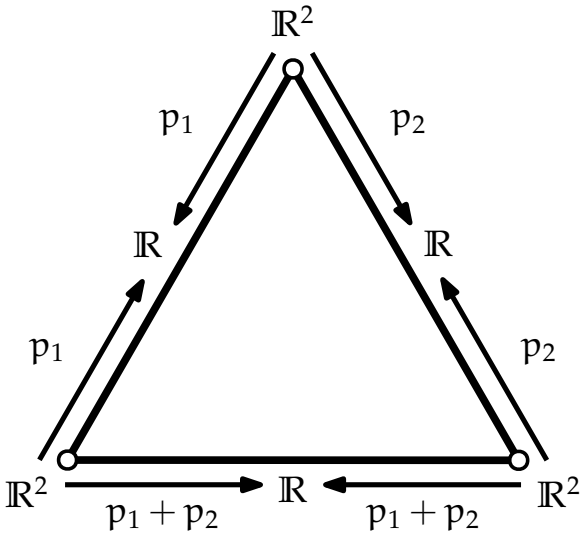


Figure 5.1: Not an approximation to the constant sheaf.

However, this condition is not sufficient. Consider the sheaf in Figure 5.1. For any cutset of the graph, $\bigcap K_e = 0$. But $\dim C^1(G;\mathcal{F}) = 3$ while $\dim C^0(G;\mathcal{F}) = 6$, so $H^0(G;\mathcal{F})$ must have dimension at least 3, larger than $H^0(G;\underline{\mathbb{R}^2}) \simeq \mathbb{R}^2$.

Applying Proposition 5.2.1 to the cutset of a graph given by $C_v = \{e : v \trianglelefteq e\}$ shows that an approximation to the constant sheaf must have $H^0(G, G \setminus \{v\};\mathcal{F}) = 0$. Conversely, the condition $H^0(G, v;\mathcal{F}) = 0$ is a sufficient condition, as can be seen from Proposition 5.1.1. The subcomplex Y is the single vertex v. The map $H^0(G;\underline{\mathbb{V}}) \to H^0(v;\underline{\mathbb{V}})$ is clearly an isomorphism, and the map $a_v : \mathbb{V} \to \mathcal{F}(v)$ is an isomorphism, so a is an approximation on v. Since $H^0(X, v;\underline{\mathbb{V}}) = 0$, the map $H^0(X, v;\underline{\mathbb{V}}) \to H^0(X, v;\mathcal{F})$ is an isomorphism if and only if $H^0(X, v;\mathcal{F}) = 0$.

Suppose we have two graphs G and H carrying approximations $\mathcal{F}$ and $\mathcal{G}$ to the constant sheaf $\underline{\mathbb{V}}$. These approximations can be glued together to produce an approximation to the constant sheaf on a graph $G + H$ with vertex set the union of the vertex sets of G and H. All that is necessary is to add edges between vertices of G and vertices of H, assigning them subspaces K_e of $\mathbb{V}$ so that their intersection is trivial. Any section of the resulting sheaf must restrict to a section of $\mathcal{F}$ on G and a section of $\mathcal{G}$ on H, so it is constant on these two subgraphs; the only way it can fail to be constant is if it has different values on each subgraph, which is prevented by the subspace intersection hypothesis.

5.3 SPARSIFICATION

Sparsification is a natural operation to apply to graphs and complexes. Since Spielman and Teng's work on spectral approximation of graphs [ST11], much of the focus has been on producing *spectrally good* sparsifications of graphs.

A sparse approximation G' of a given graph G should share properties in common with its dense relative. One particularly important property is the size of cuts. We might like, for any set of vertices A, the (weighted) number of edges between A and its complement to be similar in both G and G'. Writing this in terms of the Laplacians of G and G' says that we want $\langle \mathbb{1}_A, L_G \mathbb{1}_A \rangle \approx \langle \mathbb{1}_A, L_{G'} \mathbb{1}_A \rangle$. But why restrict to indicator vectors only? Why not simply require that $\langle x, L_G x \rangle \approx \langle x, L_{G'} x \rangle$ for all x? This is the notion of spectral approximation, and it is formalized (sometimes with slight variations) in the following definition:

Definition 5.3.1. Let G be a weighted graph. A graph G' with the same vertex set is an ϵ-spectral approximation of G if for every $x \in C^0(G; \mathbb{R})$,

$$(1-\epsilon)\langle x, L_G x \rangle \leqslant \langle x, L_{G'} x \rangle \leqslant (1+\epsilon)\langle x, L_G x \rangle.$$

This requirement can be more economically described with the standard partial order on the cone of positive semidefinite matrices. Given two PSD matrices A and B, we say that $A \preceq B$ if $B - A$ is positive semidefinite. (This order is sometimes known as the Loewner order on the cone of PSD matrices.) The requirement for G' to ϵ-approximate G is then

$$(1-\epsilon)L_G \preceq L_{G'} \preceq (1+\epsilon)L_G.$$

Spielman and Teng [ST11] first gave an algorithm that produced an ϵ-approximator of a graph with n vertices with $O(\epsilon^{-2} n \log^c n)$ edges, which was quickly improved by Spielman and Srivastava [SS11] to $O(\epsilon^{-2} n \log n)$ edges. A later algorithm, developed in collaboration with Batson, yields ϵ-sparsifiers with $O(\epsilon^{-2} n)$ edges [BSS12]. A similar notion of spectral approximation for simplicial complexes was introduced by Osting et al. [OPW19]. Chung and Zhao, and later Kyng et al., considered the question of spectral sparsification of graphs where the relevant matrix was a connection Laplacian or a block norm diagonally dominant matrix [CZ12; Kyn+16]. All of these results may be seen as special cases of the approach taken by Silva et al., who considered the question of sparsifying arbitrary sums of positive semidefinite matrices [CSHS16].

This last approach also applies to sparsifying sheaves over cell complexes. Given a $(d+1)$-dimensional complex X with a sheaf $\mathcal{F}$, we wish to produce a complex X' with a small number of $(d+1)$-cells and a sheaf $\mathcal{F}'$ on X' such that $(1-\epsilon)(\Delta_+^d)_{\mathcal{F}} \preceq (\Delta_+^d)_{\mathcal{F}'} \preceq (1+\epsilon)(\Delta_+^d)_{\mathcal{F}}$. (We control the spectrum of only the up-Laplacian because

we are only removing top-dimensional cells, which does not affect Δ^d_-.) This can be accomplished in various ways. Perhaps the most straightforward is to apply the main theorem of Silva et al. [CSHS16]:

Theorem 5.3.1 (Silva et al.). *Let $A_1, \dots, A_m$ be $n \times n$ symmetric positive semidefinite matrices, with $A = \sum_i A_i$. For any $\epsilon \in (0,1)$ there exists a nonnegative weighting $w \in \mathbb{R}^m$ with $O(n/\epsilon^2)$ nonzero entries such that*

$$A \preceq \sum_{i=1}^{m} w_i A_i \preceq (1+\epsilon)A.$$

To use this for sheaves, we consider the decomposition

$$(\Delta^d_+)_{\mathcal{F}} = \sum_{\dim \sigma = d+1} \delta^*_\sigma \delta_\sigma$$

These are matrices of dimension $n = \dim C^d(X;\mathcal{F})$, and we can therefore find a reweighting that uses only $O(n/\epsilon^2)$ of them while ϵ-approximating $(\Delta^d_+)_{\mathcal{F}}$. This amounts to deleting the $(d+1)$-cells which receive zero weighting, while reweighting the restriction maps incident to the remaining cells by the corresponding entries of the vector w. Note that for $\epsilon < 1$ this preserves the kernel of $(\Delta^d_+)_{\mathcal{F}}$ exactly.

However, this focus on simply sparsifying the complex underlying a sheaf is insufficiently ambitious. A spectrally good sparsification of a sheaf $\mathcal{F}$ on a $(d+1)$-dimensional complex gives us a d-approximation $\mathcal{G}$ of $\mathcal{F}$ by simply setting the stalks over removed $(d+1)$-cells to zero. The condition that $H^d(X;\mathcal{G}) \simeq H^d(X;\mathcal{F})$ is ensured by the fact that $(\Delta^-_{\mathcal{F}})_d = (\Delta^-_{\mathcal{G}})_d$ and $\ker(\Delta^+_{\mathcal{F}})_d = \ker(\Delta^+_{\mathcal{G}})_d$. The full generalization of sparsification to cellular sheaves is not simply finding a sparse complex on which a good approximation can be constructed, but finding a spectrally good d-approximation to $\mathcal{F}$ with small $\dim C^{d+1}(X;\mathcal{G})$.

5.3.1 *Sheaf approximation by effective resistance*

Simple extensions to the two methods for graph and matrix sum sparsification suffice to produce sheaf approximations. These extensions amount to making one additional choice at each step in the sparsification procedure.

We give first an approach based on sheaf effective resistance, following Spielman and Srivastava. Suppose we wish to compute a d-approximation $\mathcal{G}$ of a sheaf $\mathcal{F}$ on a $(d+1)$-dimensional complex X. We will do this by spectrally approximating Δ^d_+, as this will determine restriction maps for $\mathcal{G}$ between d- and $(d+1)$-cells.

Consider the matrix $\Pi = \delta^d(\Delta_+^d)^\dagger(\delta^d)^*$. This is an orthogonal projection matrix acting on $C^{d+1}(X;\mathcal{F})$, with image $\operatorname{im}\delta^d$. If S is a block diagonal symmetric positive semidefinite matrix with

$$\|\Pi S\Pi - \Pi\|_2 \leqslant \epsilon, \tag{5.1}$$

then $(\delta^d)^* S \delta^{d-1}$ ϵ-approximates Δ_+^d. This holds because the norm inequality is equivalent to

$$\frac{|\langle \Pi y, (S-I)\Pi y\rangle|}{\|y\|^2} \leqslant \epsilon$$

for every $y \in C^{d+1}(X;\mathcal{F})$. Since Π is the projection onto $\operatorname{im}\delta^d$, this implies that for $y \in \operatorname{im}\delta^d$,

$$\frac{|\langle y, (S-I)y\rangle|}{\|y\|^2} \leqslant \epsilon.$$

We reparameterize $y \in \operatorname{im}\delta^d$ by $y = \delta^d x$ for $x \perp \ker\delta^d$, so that (dropping the dimensions for notational concision) we have

$$\frac{\left|\langle \delta x, S\delta x\rangle - \langle x, \Delta_+ x\rangle\right|}{\langle x, \Delta_+ x\rangle} \leqslant \epsilon$$

for all $x \perp \ker\delta$. Thus, for instance, $\langle \delta x, S\delta x\rangle \leqslant (1+\epsilon)\langle x, \Delta_+ x\rangle$ for all x, since for $x \in \ker\delta$ this is an equality.

We will now describe a method of choosing a random matrix S so that the inequality (5.1) holds in expectation, and hence is satisfied by some realization of S. This construction relies on a matrix concentration bound as follows:

Theorem 5.3.2 (Rudelson-Vershynin [RV07]). *Let $y_1, \dots, y_q$ be independent identically distributed samples from a random distribution such that $\sup\|y_i\|_2 \leqslant M$ and $\|\mathbb{E}yy^T\| \leqslant 1$. Then*

$$\mathbb{E}\left\|\frac{1}{q}\sum_{i=1}^{q} y_i y_i^T - \mathbb{E}yy^T\right\|_2 \leqslant \min\left(CM\sqrt{\frac{\log q}{q}}, 1\right)$$

for some C independent of M and q.

We will choose our random variable y so that $\mathbb{E}yy^T = \Pi$, and take sufficiently many samples so that the norm bound is less than ϵ. Without loss of generality, we will assume that for each $(d+1)$-cell σ of X, the coboundary map δ is surjective onto σ. This ensures that $R_{\text{eff}}(\sigma) = \delta_\sigma(\Delta_+^d)^\dagger_{\mathcal{F}}\delta_{\sigma^*}$ is nonsingular.

The variable y_i will correspond to a randomly chosen $(d+1)$-cell σ_i, with probability proportional to $\operatorname{tr} R_{\mathrm{eff}}(\sigma)$. Since $R_{\mathrm{eff}}(\sigma)$ is the diagonal (σ,σ) block of Π, $\sum_\sigma \operatorname{tr} R_{\mathrm{eff}}(\sigma) = \operatorname{tr}\Pi = \dim\operatorname{im}\delta = \operatorname{codim}\ker\delta$. Therefore, the probability of choosing σ is $p_\sigma = \frac{\operatorname{tr} R_{\mathrm{eff}}(\sigma)}{\dim\operatorname{im}\delta}$.

Once we have chosen the cell σ, we choose a random vector $x_\sigma \in \mathcal{F}(\sigma)$ so that $\mathbb{E}x_\sigma x_\sigma^T = \mathrm{id}_{\mathcal{F}(\sigma)}$ and $x_\sigma^T R_{\mathrm{eff}}(\sigma) x_\sigma = \operatorname{tr} R_{\mathrm{eff}}(\sigma)$. In particular, letting v_σ^i be the unit eigenvectors of $R_{\mathrm{eff}}(\sigma)$, with eigenvalues λ_i we let $x_\sigma = \sqrt{\frac{\operatorname{tr} R_{\mathrm{eff}}(\sigma)}{\lambda_i}} v_\sigma^i$. Then

$$\mathbb{E}xx^T = \sum_i \frac{\lambda_i}{\operatorname{tr} R_{\mathrm{eff}}(\sigma)} \frac{\operatorname{tr} R_{\mathrm{eff}}(\sigma)}{\lambda_i} v_i v_i^T = \mathrm{id}_{\mathcal{F}(\sigma)}$$

and

$$x_i^T R_{\mathrm{eff}}(\sigma) x_i = \frac{\operatorname{tr} R_{\mathrm{eff}}(\sigma)}{\lambda_i} v_i^T R_{\mathrm{eff}}(\sigma) v_i = \frac{\operatorname{tr} R_{\mathrm{eff}}(\sigma)}{\lambda_i} \lambda_i = \operatorname{tr} R_{\mathrm{eff}}(\sigma).$$

After choosing σ and x_σ, we let $y = \frac{1}{\sqrt{p_\sigma}} \Pi x_\sigma$.

We now need to check that $\mathbb{E}yy^T = \Pi$, and bound $\|y\|_2$. We have

$$\mathbb{E}yy^T = \sum_\sigma p_\sigma \frac{1}{p_\sigma} \mathbb{E}\Pi x_\sigma x_\sigma^T \Pi = \sum_\sigma \Pi \,\mathrm{id}_{\mathcal{F}(\sigma)}\, \Pi = \Pi.$$

Since $y = \frac{1}{\sqrt{p_\sigma}} \Pi x_\sigma$ for some σ, we can bound $\|y\|$ by bounding

$$\left\| \frac{1}{\sqrt{p_\sigma}} \Pi x_\sigma \right\|^2 = \frac{1}{p_\sigma} x_\sigma^T \Pi x_\sigma = \frac{1}{p_\sigma} x_\sigma^T R_{\mathrm{eff}}(\sigma) x_\sigma = \frac{\dim\operatorname{im}\delta}{\operatorname{tr} R_{\mathrm{eff}}(\sigma)} \operatorname{tr} R_{\mathrm{eff}}(\sigma) = \dim\operatorname{im}\delta.$$

Thus $\|y\| \leqslant \sqrt{\dim(\operatorname{im}\delta)}$. Let $B = \dim(\operatorname{im}\delta))$. We then have

$$\mathbb{E}\|\sum_{i=1}^q y_i y_i^T - \Pi\|_2 \leqslant C\sqrt{B}\sqrt{\frac{\log q}{q}}.$$

Note that the matrix S is equal to $\sum_i \frac{1}{p_{\sigma_i}} x_{\sigma_i}^i (x_{\sigma_i}^i)^T$, which implies an approximation morphism $a_\sigma : \mathcal{F}(\sigma) \to \mathcal{G}(\sigma)$ given by the projection onto the vectors $\frac{1}{\sqrt{p_\sigma}} x_\sigma^i$ for the various realizations of x_σ used in the construction.

What remains is to choose q, which controls the total dimension of $C^{d+1}(X;\mathcal{G})$. We want the bound to be at most ϵ, so we need

$$\frac{q}{\log q} \geqslant \frac{C^2 B}{\epsilon^2}.$$

Choosing $q = \frac{tC^2B\log(B)}{\epsilon^2}$ gives us

$$\frac{q}{\log q} = \frac{tC^2B\log B}{\epsilon^2 \log\left(\frac{2C^2B\log B}{\epsilon^2}\right)}.$$

If B is sufficiently large that

$$(B)^t \geqslant \frac{C^2B\log B}{\epsilon^2},$$

then we have

$$\begin{aligned} t\log B &\geqslant \log\left(\frac{C^2B\log B}{\epsilon^2}\right) \\ \frac{tC^2B\log B}{\epsilon^2} &\geqslant \frac{C^2B}{\epsilon^2}\log\left(\frac{C^2B\log B}{\epsilon^2}\right) \\ \frac{tC^2B\log B}{\epsilon^2\log\left(\frac{2C^2B\log B}{\epsilon^2}\right)} &\geqslant \frac{C^2B}{\epsilon^2}. \end{aligned}$$

This discussion has closely followed the original proof of the spectral sparsification algorithm by Spielman and Srivastava [SS11]. The eigendecomposition of $R_{\text{eff}}(\sigma)$ can be seen as giving a ranking of subspaces of $\mathcal{F}(\sigma)$ by their importance to the spectrum of $(\Delta_+^d)_{\mathcal{F}}$.

In summary, we have shown

Theorem 5.3.3. *Let $\mathcal{F}$ be a sheaf on a $(d+1)$-dimensional cell complex X. For any $\epsilon \in (0,1)$ There exists a d-approximation $\mathcal{G}$ to $\mathcal{F}$ such that $(1-\epsilon)(\Delta_+^d)_{\mathcal{F}} \preceq (\Delta_+^d)_{\mathcal{G}} \preceq (1+\epsilon)(\Delta_+^d)_{\mathcal{F}}$ and $\dim C^{d+1}(X;\mathcal{G}) = O(\dim(\operatorname{im}\delta_{\mathcal{F}}^d)\log(\dim(\operatorname{im}\delta_{\mathcal{F}}^d))\epsilon^{-2})$.*

Note that $\dim(\operatorname{im}\delta^d) \leqslant \dim C^d(X;\mathcal{F})$, so the dimension of $C^{d+1}(X;\mathcal{F}')$ is in fact bounded in terms of the dimension of C^d.

One could get an asymptotically similar bound by simply choosing beforehand a decomposition of each matrix $\delta_\sigma^*\delta_\sigma$ into a sum of rank-1 matrices and applying the $O(n\log n/\epsilon^2)$ approximation algorithm of Silva et al. [CSHS16] This proof illustrates the preferability of choosing a particular rank-1 decomposition of these matrices, namely that given by the eigenvectors of the effective resistance matrix for each $(d+1)$-cell.

There is a natural randomized algorithm to produce such an approximation. It requires computation of Π, or at least the effective resistances $R_{\text{eff}}(\sigma)$ for all $(d+1)$-cells σ. Naively, this takes $O((\dim C^d(X;\mathcal{F}))^3)$ time, by finding the pseudoinverse of $(\Delta_+^d)_{\mathcal{F}}$. However, the special structure of the up-Laplacian may allow for quicker computations, or at least approximations, of the effective resistances. Once the

effective resistances are known, the algorithm takes $O(q)$ steps which involve sampling vectors of size $O(D)$.

5.3.2 *Linear-sized approximations*

We can also produce sheaf approximations by generalizing the approach used in [BSS12; CSHS16] to produce linear-sized graph or matrix sum sparsifiers. The algorithms to produce these are not randomized, but are still nondeterministic. The main approach is to iteratively select matrices to add to the sum while maintaining careful bounds on how well the partial sum approximates the desired matrix.

We will begin by converting the problem into that of approximating a sum of positive semidefinite matrices. We want to find a sheaf $\mathcal{G}$ with small $\dim C^{d+1}(X;\mathcal{G})$ such that $(\Delta_+^d)_\mathcal{G}$ is a good spectral approximation of $(\Delta_+^d)_\mathcal{F}$. Write

$$(\Delta_+^d)_\mathcal{F} = \sum_{\dim\sigma=d+1} (\delta_\mathcal{F}^d)_\sigma^*(\delta_\mathcal{F}^d)_\sigma = \sum_{\dim\sigma=d+1} \Delta_\mathcal{F}^\sigma$$

and

$$(\Delta_+^d)_\mathcal{G} = \sum_{\dim\sigma=d+1} (\delta_\mathcal{G}^d)_\sigma^*(\delta_\mathcal{G}^d)_\sigma = \sum_{\dim\sigma=d+1} \Delta_\mathcal{G}^\sigma.$$

We want to be able to write $(\delta_\mathcal{G}^d)_\sigma = a_\sigma(\delta_\mathcal{F}^d)_\sigma$ for some collection of maps a_σ defining an approximation $a : \mathcal{F} \to \mathcal{G}$. It is possible to do this if $\ker\Delta_\mathcal{G}^\sigma \supseteq \ker\Delta_\mathcal{F}^\sigma$ for all σ.

Lemma 5.3.1. *Let* $\ker\Delta_\mathcal{G}^\sigma \supseteq \ker\Delta_\mathcal{F}^\sigma$ *for all* σ. *Then* $\sum_\sigma \Delta_\mathcal{G}^\sigma$ *is the degree-*d *up-Laplacian of a sheaf* $\mathcal{G}$ *carrying a morphism* $a : \mathcal{F} \to \mathcal{G}$.

Proof. Let $\mathcal{G}(\sigma) = \mathcal{F}(\sigma)$ for all σ of dimension at most d, and let $a_\sigma = \mathrm{id}$ for these cells. Take an orthogonal basis of eigenvectors of $\Delta_\mathcal{G}^\sigma$ which contains a basis for $\ker\Delta_\mathcal{F}^\sigma$. There exists a matrix A_σ so that for every v, w in this basis, we have $\langle(\delta_\mathcal{F}^d)_\sigma v, A_\sigma(\delta_\mathcal{F}^d)_\sigma w\rangle = \langle v, \Delta_\mathcal{G}^\sigma w\rangle$, ensuring that $(\delta_\mathcal{F}^d)_\sigma^* A_\sigma(\delta_\mathcal{F}^d)_\sigma = \Delta_\mathcal{G}^\sigma$. A_σ is constructed by noting that for eigenvectors $v \perp \ker\Delta_\mathcal{F}^\sigma$, $(\delta_\mathcal{F}^d)_\sigma v$ form a basis for $\mathrm{im}(\delta_\mathcal{F}^d)_\sigma$, so an operator on this space is determined by the inner products above. Taking a rank-revealing decomposition $A_\sigma = a_\sigma^* a_\sigma$ gives us stalks $\mathcal{G}(\sigma)$ for $(d+1)$-cells σ, as well as the desired approximation maps $a_\sigma : \mathcal{F}(\sigma) \to \mathcal{G}(\sigma)$.

Once we have the maps a_σ, the restriction maps $\mathcal{G}_{\tau\trianglelefteq\sigma}$ are readily constructed by requiring these stalkwise maps to form a sheaf morphism. These restriction maps determine a functor $\mathcal{G} : X \to \mathbf{Hilb}$ by a quick diagram chase given the commutativity of a with the restriction maps. □

Thus, we will be able to compute spectrally good d-approximations to $\mathcal{F}$ if the following holds:

Theorem 5.3.4. *Let $\{A_i\}$ be a collection of positive semidefinite $D \times D$ matrices. For any $\epsilon \in (0,1)$ there exists a collection of positive semidefinite matrices $\{B_i\}$ with $\ker B_i \supseteq \ker A_i$ and $\sum_i \operatorname{rank}(B_i) = O(D/\epsilon^2)$ such that*

$$(1-\epsilon)\sum_i A_i \preceq \sum_i B_i \preceq (1+\epsilon)\sum_i A_i.$$

We will prove this theorem by an extension of de Carli Silva et al.'s proof of Theorem 5.3.1, which is itself an extension of Batson, Spielman, and Srivastava's proof of the existence of linear-sized graph sparsifiers.

We first reduce to the case where $\sum_i A_i = I$, which is clearly possible by restricting to $\ker(\sum_i A_i)^\perp$ and multiplying by $(\sum_i A_i)^{-1/2}$ on either side. This allows us to deduce the Loewner order inequality by controlling the largest and smallest eigenvalues of $\sum_i B_i$. The proof proceeds by iteratively selecting rank-1 matrices C_k, with $\ker C_k \supseteq \ker A_{i_k}$ for some i_k, to add to a running sum, with bounds on the upper and lower eigenvalues of the sum carefully maintained. These bounds are ensured by a pair of "barrier functions" whose values control the change in eigenvalues as we add a new matrix C_k. It will then be possible to show that after we add $O(D/\epsilon^2)$ such matrices, the approximation bound will be satisfied.

For $u > \lambda_{\max}(A)$, $\Phi^u(A) = \operatorname{tr}(uI - A)^{-1}$ is finite. As u approaches $\lambda_{\max}(A)$, $\Phi^u(A)$ goes to infinity. Similarly, for $\ell < \lambda_{\min}(A)$, $\Phi_\ell = \operatorname{tr}(A - \ell I)^{-1}$ is finite, and goes to infinity as ℓ approaches $\lambda_{\min}(A)$. The goal is to increment u and ℓ appropriately as we add new terms to our matrix A, so that the eigenvalues of A never pass below ℓ or above u, and after a predetermined number of steps A will be close to the identity.

Our goal at each step is then to choose a matrix C so that $\lambda_{\max}(A+C)$ and $\Phi^{u+\delta_u}(A+C)$ are controlled, and similarly for $\lambda_{\min}(A+C)$ and $\Phi^{\ell+\delta_\ell}(A+C)$. Conditions for this to be possible are given by a pair of lemmas proved by Batson, Spielman, and Srivastava [BSS12] and generalized by de Carli Silva et al. [CSHS16]. We will let $A^u = (uI - A)^{-1}$ and $A_\ell = (A - \ell I)^{-1}$.

Lemma 5.3.2 (Lemmas 4.1 and 4.2, [CSHS16])**.** *Let A be a symmetric matrix, and C symmetric positive semidefinite and nonzero; let $\ell, u \in \mathbb{R}$, and $\delta_\ell, \delta_U > 0$. If $\ell < \lambda_{\min}(A) \leqslant \lambda_{\max}(A) < u$ and*

$$\frac{\langle (A^{u+\delta_u})^2, X\rangle}{\Phi^u(A) - \Phi^{u+\delta_u}(A)} + \langle A^{u+\delta_u}, X\rangle \leqslant \frac{1}{\alpha} \leqslant \frac{\langle A^2_{\ell+\delta_\ell}, X\rangle}{\Phi_{\ell+\delta_\ell}(A) - \Phi_\ell(A)} - \langle A_{\ell+\delta_\ell}, X\rangle, \quad (5.2)$$

then $\ell + \delta_\ell < \lambda_{\min}(A + \alpha X) \leqslant \lambda_{\max}(A + \alpha X) < u + \delta_u$, $\Phi^{u+\delta_u}(A + \alpha X) \leqslant \Phi^u(A)$, *and* $\Phi_{\ell+\delta_\ell}(A + \alpha X) \leqslant \Phi_\ell(A)$.

If A is our running sum, we need to choose the matrix X from the set of admissible matrices: those that have kernel containing $\ker A_i$ for some A_i. The goal is to choose appropriate values of δ_ℓ, δ_u, and bounds on Φ^u and Φ_ℓ so that it is always possible to do this.

The necessary result is this:

Lemma 5.3.3 (Lemma 4.3, [CSHS16])**.** *Suppose* $\ell < \lambda_{\min}(A) \leqslant \lambda_{\max}(A) < u$, $\Phi^u(A) \leqslant \epsilon_U$, *and* $\Phi_\ell(A) \leqslant \epsilon_L$. *If*

$$0 \leqslant \frac{1}{\delta_u} + \epsilon_U \leqslant \frac{1}{\delta_\ell} - \epsilon_L,$$

then there exists some j *and* α *such that* (5.2) *is satisfied.*

The proof proceeds by showing that there exists some α such that the inequality holds in expectation for A_i chosen uniformly at random, which implies that it holds for some A_i.

We know, therefore, that at each step k in constructing A, there exists some A_{i_k} such that we can add αA_{i_k} to the sum and preserve the necessary bounds. But we can go further: once we know A_{i_k}, we can choose a random vector c such that $\mathbb{E}[cc^*] = \alpha A_{i_k}$ and $c \perp \ker A_{i_k}$. Then by linearity, the inequality (5.2) holds in expectation for the matrix cc^*, and hence there exists some particular realization c_k such that it holds. We then use $C_k = \alpha' c_k c_k^*$ as the next term in the sum, for an appropriate α'.

What remains is to choose the parameters appropriately. Given our approximation level ϵ, we let

$$\delta_L = 1, \quad \epsilon_L = \frac{\epsilon}{2}, \quad \ell_0 = -\frac{D}{\epsilon_L}, \quad \delta_U = \frac{2+\epsilon}{2-\epsilon}, \quad \epsilon_U = \frac{\epsilon}{2\delta_U}, \quad u_0 = \frac{D}{\epsilon_U},$$

again as previously derived. After $T = \frac{4D}{\epsilon^2}$ steps, we have $\lambda_{\max}(A) \leqslant u_0 + T\delta_U = \left(\frac{2D}{\epsilon} + \frac{4D}{\epsilon^2}\right)\frac{2+\epsilon}{2-\epsilon}$ and $\lambda_{\min}(A) \geqslant \ell_0 + T\delta_L = -\frac{2D}{\epsilon} + \frac{4D}{\epsilon^2}$. Thus

$$\frac{\lambda_{\max}(A)}{\lambda_{\min}(A)} \leqslant \frac{u_0 + T\delta_U}{\ell_0 + T\delta_L} = \frac{1 + \frac{2}{\epsilon}}{-1 + \frac{2}{\epsilon}} \frac{2+\epsilon}{2-\epsilon} = \left(\frac{2+\epsilon}{2-\epsilon}\right)^2 \leqslant \frac{1+\epsilon}{1-\epsilon}.$$

This ensures that

$$(1-\epsilon)I \preceq \frac{1-\epsilon}{\lambda_{\min}(A)} A \preceq (1+\epsilon)I,$$

which is our desired bound. Further, A is the sum of T rank-one matrices C_k, and hence when we collate these matrices together into $B_i = \sum_{k:i_k=i} C_k$, $\sum_i \operatorname{rank}(B_i) \leqslant T = \frac{4D}{\epsilon^2}$. We in fact have a concrete bound, not just an asymptotic one.

Ascending from these details back to the question of sheaf approximation, we have shown (with assistance from Batson et al. and de Carli Silva et al.):

Theorem 5.3.5. *Let $\mathcal{F}$ be a sheaf on a $(d+1)$-dimensional cell complex X. For any $\epsilon \in (0,1)$, there exists a d-approximation $\mathcal{G}$ of $\mathcal{F}$ with $(1-\epsilon)(\Delta_+^d)_{\mathcal{F}} \preceq (\Delta_+^d)_{\mathcal{G}} \preceq (1+\epsilon)(\Delta_+^d)_{\mathcal{F}}$ and $\dim C^{d+1}(X;\mathcal{G}) = 4\dim(C^d(X;\mathcal{F}))/\epsilon^2$.*

Neither this algorithm nor the effective resistance-based algorithm is able to give any guarantees about the dimension of a particular stalk in the approximating sheaf. If a particular $(d+1)$-cell has relatively large effective resistance, it may be chosen sufficiently often that its stalk dimension is entirely preserved. Indeed, it is not in general possible to produce a d-approximation to a sheaf with a uniform reduction in stalk dimensions. Consider, for instance, the constant sheaf $\underline{\mathbb{R}^n}$ on a graph with two densely connected subgraphs joined by a single edge e. Any approximation $\mathcal{F}$ must have $\mathcal{F}(e) = \mathbb{R}^n$ by Proposition 5.2.1. Determining conditions under which it is possible to uniformly reduce stalk dimensions in an approximation will require a more subtle analysis.

6

EXPANSION

6.1 EXPANDER GRAPHS

In graphs, "expansion" refers to certain qualitative measures of connectedness. These properties were initially defined combinatorially, but turn out to have useful and interesting spectral analogues. Intuitively, a graph is well connected if it is hard to disconnect: removing a few edges or vertices should leave us with a connected graph. This gives the combinatorial notion of vertex or edge connectivity of a graph—one less than the minimal number of vertex or edge deletions necessary to disconnect the graph. Expansion is a somewhat subtler concept, depending not only on the amount of work necessary to disconnect the graph, but also on the size of the resulting components.

Definition 6.1.1. Let G be a graph. The *edge expansion constant* of G is

$$h(G) = \min_{A \subseteq V(G)} \frac{|\partial A|}{\min(|A|, |V(G) \setminus A|)}.$$

The *vertex expansion constant* of G is

$$h_v(G) = \min_{A \subseteq V(G)} \frac{|N(A)| - |A|}{\min(|A|, |V(G) \setminus A|)}.$$

The term "expansion" comes from thinking of the ratio of the size of the vertex set A to the size of its neighborhood. If the vertex expansion constant of G is large, every sufficiently small set of vertices has a neighborhood that expands its size significantly. To disconnect a set A of size less than half the number of vertices of the graph, one must sever at least $h(G)|A|$ edges, or delete at least $h_v(G)|A|$ vertices.

Another common measure of connectivity of a graph is the *algebraic connectivity*, the second smallest eigenvalue of the graph Laplacian. By the Courant-Fischer-Weyl theorem, this is

$$\lambda_2 = \min_{x \perp \mathbb{1}} \frac{\langle x, Lx \rangle}{\langle x, x \rangle}.$$

There are various results relating the algebraic connectivity of a graph with its vertex and edge expansion. Perhaps the most famous is the Cheeger inequality, so called by analogy with Cheeger's isoperimetric inequality for Riemannian manifolds.

Theorem 6.1.1 (Cheeger inequality, [Dod84])**.** *Let* G *be a* d*-regular graph, with algebraic connectivity* λ_2*. Then*

$$\frac{\lambda_2}{2} \leqslant h(G) \leqslant \sqrt{2d\lambda_2}.$$

The left-hand side of this inequality is straightforward to prove, since the minimax characterization of λ_2 can be seen as a relaxation of the optimization problem defining $h(G)$. The right-hand side is more involved, but gives a construction of a set of vertices with edge expansion at most $\sqrt{2d\lambda_2}$ from the corresponding eigenvector of L. Stronger versions of this inequality are in fact true, but this version suffices to show an equivalence between combinatorial and spectral measures of expansion.

For regular graphs, the spectra of the adjacency matrix and the Laplacian determine each other—an eigenvalue λ of L gives an eigenvalue $\mu = d - \lambda$ of A. Thus we can cast spectral expansion in terms of eigenvalues of the adjacency matrix, which introduces a new spectral property of interest. The algebraic connectivity λ_2 becomes an eigenvalue $\mu_2 = d - \lambda_2$, while the eigenvalue $\lambda_1 = 0$ becomes the eigenvalue $\mu_1 = d$. It is straightforward to see (for instance, by applying Proposition 4.1.3) that the maximum eigenvalue of L is at most $2d$, so the minimum eigenvalue of A is at least $-d$. We can thus either put a symmetric bound on the nontrivial eigenvalues of A, requiring them to lie close to zero, or only put an upper bound, requiring them to stay away from d.

Definition 6.1.2. A d-regular graph G has *two-sided spectral expansion constant* η if all its nontrivial adjacency eigenvalues (all but the eigenvalue d) lie in the interval $[-d+\eta, d-\eta]$. G has *one-sided* spectral expansion constant η if all nontrivial adjacency eigenvalues lie in $[-d, d-\eta]$.

A graph with an expansion constant bounded by η (either above or below, as appropriate for combinatorial or spectral expansion constants) is called an η-expander. One way to interpret the Cheeger inequality is that being an edge expander and being a one-sided spectral expander are equivalent up to a factor of a square root.

A surprising result due to Pinsker[1] [Pin73] (to whom we also owe the term "expander") is

1 It appears a similar result was generated by Kolmogorov and Barzdin several years earlier in a different context [KB67].

Theorem 6.1.2 (Pinsker). *Let $G_{n,d}$ be a uniform random d-regular graph on n vertices, for $d \geqslant 3$. There exists some $\epsilon > 0$ such that $\mathbb{P}(h(G_{n,d}) \geqslant \epsilon) \to 1$ as $n \to \infty$. Therefore, there is an infinite family of d-regular ϵ-edge expander graphs.*

This result is somewhat surprising. In essence, it is possible to maintain a uniform level of connectivity in a family of graphs, even as the number of edges grows only linearly with the number of vertices.

Applying the Cheeger inequality shows that there also exists an infinite family of d-regular one-sided spectral expanders for some nontrivial spectral expansion constant.

Once the existence of infinite expander families has been established, a natural question is then how good can expanders be? This is hard to answer for combinatorial expansion constants, but for spectral expansion a bound on η was established by Alon and Boppana[2]:

Theorem 6.1.3 (Alon-Boppana). *The largest nontrivial adjacency eigenvalue of a d-regular graph is asymptotically at least $2\sqrt{d-1} - o(1)$ as the number of vertices goes to infinity.*

This puts a limit on precisely how good a spectral expander can be. There can be no infinite family of d-regular expanders with expansion constant better than $\eta = d - 2\sqrt{d-1}$. We then have the question of whether this bound can be achieved. A d-regular graph satisfying the eigenvalue bound $\mu_2 \leqslant 2\sqrt{d-1}$ is called a *Ramanujan graph*. The first constructions of infinite families of Ramanujan graphs are due to, independently, Margulis and Lubotzky-Phillips-Sarnak [Mar88; LPS88].

Theorem 6.1.4 (Margulis, Lubotzky-Phillips-Sarnak). *For any prime p, there exists an infinite family of $(p^k + 1)$-regular Ramanujan graphs, realized as Cayley graphs of $\mathrm{PSL}(2, \mathbb{F}_q)$ for an appropriately chosen q.*

This was recently strengthened by Marcus, Spielman, and Srivastava [MSS15; MSS18] to

Theorem 6.1.5 (Marcus-Spielman-Srivastava). *For every $d \geqslant 3$, there exist d-regular (bipartite) Ramanujan graphs with every number of vertices.*

Note that bipartiteness implies that these graphs have $-d$ as an adjacency eigenvalue and hence are not technically double-sided spectral expanders.

2 This theorem is credited to Alon and Boppana, but it seems that they never published a proof. A well known proof was published by Nilli in [Nil91].

6.2 EXPANDER SHEAVES

Our goal in this section is to develop an analogue of expander graphs in the context of cellular sheaves. Although some study has been made of the concept of expansion in higher-dimensional simplicial complexes [Gro09; Gro10; Par13; Ste13], we will limit our focus to sheaf-theoretic analogues of expander graphs, as this is a sufficiently difficult problem.

Combinatorial definitions are hard to generalize to algebraic structures like sheaves, so we will begin from the spectral perspective. An initial approach might be to simply let an expander sheaf be a "regular" sheaf on a graph with eigenvalues concentrated in some small set. However, this definition is not very interesting: there exist cellular sheaves whose Laplacian is the identity matrix, giving them a perfect "expansion constant." Another difficulty is that expander graphs are unweighted, while the restriction maps of a sheaf are inherently weighted. This, combined with a desire to maintain an analogy with the properties of expander graphs, motivates a decision to restrict to matrix-weighted graphs with very particular sorts of weights.

Definition 6.2.1. A 0-approximation $\mathcal{F}$ to $\underline{\mathbb{R}^n}$ over a graph G is an η-*expander sheaf* if all of the following hold:

1. $\mathcal{F}$ is dI-regular
2. Every restriction map is a partial isometry
3. All its nontrivial adjacency eigenvalues (i. e.those apart from the eigenvalue d with multiplicity n) lie in $[-d+\eta, d-\eta]$.

Constant sheaves over a d-regular expander graph clearly qualify as expander sheaves. Note that if G is an ℓ-regular graph and all restriction maps of an expander sheaf $\mathcal{F}$ are rank k, with vertex stalks of dimension n, then $d = \ell\frac{k}{n}$. However, it is not immediately obvious that it is possible to nontrivially satisfy these conditions, let alone with well-controlled values of η.

Theorem 6.2.1. *There exist nontrivial examples of expander sheaves.*

To show this, we will first need a couple of definitions from finite-dimensional frame theory.

Definition 6.2.2. Let $\mathcal{H}$ be a Hilbert space. A *frame* in $\mathcal{H}$ is a collection of vectors F in $\mathcal{H}$ such that there exist positive A, B such that for every $x \in \mathcal{H}$, we have

$$A\|x\|^2 \leqslant \sum_{e\in F} |\langle e, x\rangle|^2 \leqslant B\|x\|^2.$$

If we may take $A = B$, F is called a *tight frame*.

A frame is a sort of "linearly dependent basis" for $\mathcal{H}$. Frame theory arose in the study of nonunique signal decompositions such as wavelets. However, if $\mathcal{H}$ is finite-dimensional, frames (at least those with finitely many vectors) are equivalent to matrices of full rank.

For a frame F, we have the *synthesis operator* $S_F : \mathbb{R}^F \to \mathcal{H}$ given by $S_F(y) = \sum_{e \in F} y_e e$, and the *analysis operator* $A_F : \mathcal{H} \to \mathbb{R}^F$ given by $A_F(x) = (\langle e, x \rangle)_{e \in F}$. When we assign $\mathbb{R}^F$ its natural Hilbert space structure, these maps are adjoints; when F is a finite set, the matrix representing S_F has columns equal to the frame elements, and A_F has rows equal to the frame elements. We further have $\sum_{e \in F} |\langle e, x \rangle|^2 = \|A_F(x)\|^2 = \langle A_F(x), A_F(x) \rangle$. Thus, a frame is tight if and only if $A_F^* A_F = S_F A_F$ is a scalar multiple of the identity.

A frame all of whose elements have norm 1 is a *unit norm frame*. We will use a generalization of unit norm frames to construct nontrivial expander sheaves.

Definition 6.2.3. Let $\mathcal{H}$ be a Hilbert space. A *fusion frame* in $\mathcal{H}$ is a collection F of subspaces $V \leqslant \mathcal{H}$ such that if P_V is the orthogonal projection onto V, we have

$$A\|x\|^2 \leqslant \sum_{V \in F} \|P_V x\|^2 \leqslant B\|x\|^2$$

for every $x \in \mathcal{H}$.

A fusion frame whose subspaces all have dimension 1 is the same as a unit norm frame. Again, a fusion frame is *tight* if we may take $A = B$ in the above inequality. Therefore, a tight fusion frame is one for which the projection matrices P_V sum to a scalar multiple of the identity.

Nontrivial tight fusion frames exist. Specifically, Casazza et al. showed that for $d \geqslant \lceil \frac{n}{\ell} \rceil + 2$, there exists a fusion frame in $\mathbb{R}^n$ consisting of d subspaces of dimension ℓ [Cas+11]. Further, no such fusion frames exist for $d \leqslant \lceil \frac{n}{\ell} \rceil$.

We can use fusion frames to construct regular approximations to the constant sheaf on a regular graph. Specifically, let G be a d-regular graph with a d-edge coloring, and suppose we wish to approximate the constant sheaf $\underline{\mathbb{R}^n}$ on G. We choose a tight fusion frame F for $\mathbb{R}^n$ with d elements. In particular, then, we have $\sum_{V \in F} P_V = \frac{d}{n} I$. We use the projections P_V as the weights of a matrix-weighted graph. Since there are d projections for the fusion frame and d edge colors, we match each edge color to a projection and assign these projections as weights to the corresponding edges. Since each vertex is incident to precisely one edge of each color, the matrix-weighted degree for each vertex is equal to $\frac{d}{n} I$.

While this construction will produce a $\frac{d}{n} I$-regular matrix-weighted graph with orthogonal projections as weights, there is no guarantee that this will be an approximation to the constant sheaf. The resulting sheaf may have additional global sections, and hence have nontrivial adjacency eigenvalues equal to $\frac{d}{n}$. However, we

can construct individual instances where this is not the case; for instance, the sheaf described in Figure 3.1 is one such sheaf on the 3-regular graph K_4; the frame used is the *Mercedes-Benz frame*, the simplest nontrivial frame in $\mathbb{R}^2$.

The constructions of tight fusion frames by Casazza et al. are reassuring in that they imply that it is possible to construct expander-like sheaves. However, their particular constructions are suboptimal for this purpose, because many of the subspaces turn out to be spanned by standard basis vectors of $\mathbb{R}^n$. Intuitively, it seems that a fusion frame analogue of *equiangular* frames would be more desirable, since these would preserve information more equitably across the edges incident to a vertex.

6.2.1 *Bounds for expander sheaves*

The Alon-Boppana bound tells us how good a spectral expansion constant we can hope for in an infinite family of expanders. Here we search for its analogue for an expander sheaf. If we try to generalize its proof to matrix-weighted graphs, we do not obtain the same bound, but rather a scaled version. This is a general phenomenon when comparing the spectrum of a matrix-weighted graph with that of its underlying graph.

Theorem 6.2.2. *Let $\mathcal{F}$ be a matrix-weighted graph with underlying graph G and vertex stalks $\mathbb{R}^n$. Consider the weighted graph $\operatorname{tr}(\mathcal{F})$ with edge weights $w_{uv} = -\operatorname{tr}(L_{\mathcal{F}}[u,v])$; denote its Laplacian matrix L, its algebraic connectivity λ_2, and its largest Laplacian eigenvalue $\lambda_{\max}$. Let $\lambda_2(\mathcal{F})$ be the $(n+1)$st smallest eigenvalue of $L_{\mathcal{F}}$ and $\lambda_{\max}(\mathcal{F})$ the largest eigenvalue of $L_{\mathcal{F}}$. Then $\lambda_2(\mathcal{F}) \leqslant \frac{1}{n}\lambda_2$ and $\lambda_{\max}(\mathcal{F}) \geqslant \frac{1}{n}\lambda_{\max}$.*

Proof. Let y be a unit eigenvector of L corresponding to λ_2, and let x be a uniform random variable on the unit sphere in $\mathbb{R}^n$. Note that $y \otimes x$ may be seen as a 0-cochain of $\mathcal{F}$; further, $\|y \otimes x\| = 1$ and $y \otimes x$ is orthogonal to the space of constant 0-cochains of $\mathcal{F}$. We now compute

$$\begin{aligned}
\mathbb{E}[\langle y \otimes x, L_{\mathcal{F}}(y \otimes x)\rangle] &= \mathbb{E}\left[\sum_{u,v \trianglelefteq e} \langle y_v x - y_u x, W_e(y_v x - y_u x)\rangle\right] \\
&= \sum_{u,v \trianglelefteq e} (y_v - y_u)^2 \mathbb{E}\left[\langle x, W_e x\rangle\right] \\
&= \sum_{u,v \trianglelefteq e} \frac{\operatorname{tr}(W_e)}{n}(y_v - y_u)^2 \\
&= \frac{1}{n}\langle y, Ly\rangle = \frac{1}{n}\lambda_2.
\end{aligned}$$

Since this is an expectation, there must exist some unit norm x such that $\langle y \otimes x, L_{\mathcal{F}}(y \otimes x)\rangle \leqslant \frac{1}{n}\lambda_2$, and hence $\lambda_2(\mathcal{F}) \leqslant \frac{1}{n}\lambda_2$. Applying essentially the same argument with y an eigenvector of L corresponding to $\lambda_{\max}$, we get the bound $\lambda_{\max}(\mathcal{F}) \geqslant \frac{1}{n}\lambda_{\max}$. □

We apply this result to an expander sheaf $\mathcal{F}$ on a d-regular graph where all edge weights are rank k projections. Then the edge weights of $\mathrm{tr}(\mathcal{F})$ are all k, so if we take the underlying unweighted graph, we have $\lambda_2(\mathrm{tr}(\mathcal{F})) = k\lambda_2(G)$, and hence $\lambda_2(\mathcal{F}) \leqslant \frac{k}{n}\lambda_2(G)$. Similarly, $\lambda_{\max}(\mathcal{F}) \geqslant \frac{k}{n}\lambda_{\max}(G)$. The degree of $\mathcal{F}$ is $\frac{k}{n}dI$, so the nontrivial adjacency eigenvalues are at least as far from zero as $\frac{k}{n}(d-\lambda_2(G)) = \frac{k}{n}\mu_2(G)$ and $\frac{k}{n}(d-\lambda_{\max}(G)) = \frac{k}{n}\mu_N(G)$.

Now applying the Alon-Boppana bound to G, we have $|\mu(\mathcal{F})| \geqslant \frac{k}{n}|\mu(G)| \geqslant 2\frac{k}{n}\sqrt{d-1} - o(1)$. Rewriting this bound in terms of the sheaf degree $\ell = \frac{k}{n}d$, this is equal to $2\frac{\ell}{d}\sqrt{d-1} - o(1)$. In general, this weaker than the naive translation of the Alon-Boppana bound, replacing the graph degree with the sheaf degree, since $\frac{\ell}{d}\sqrt{d-1} \leqslant \sqrt{d-1}$ in general.

This suggests the exciting possibility that we may be able to construct better-than-Ramanujan expander sheaves. However, our best known techniques for constructing Ramanujan graphs do not immediately extend to sheaves, and even constructing an infinite family of nontrivial expander sheaves is difficult. On the other hand, if the direct translation of the Alon-Boppana bound does in fact hold for expander sheaves, we have a different interesting result: Ramanujan graphs are optimal solutions to the relaxed problem of finding families of sheaves with good expansion constants.

6.3 THE EXPANDER MIXING LEMMA

The expander mixing lemma for graphs, first proved by Alon and Chung [AC88], quantifies the intuitive statement that graphs with good spectral expansion are "nearly random." In a random d-regular graph, the expected number of edges between two sets S and T of vertices is

$$\mathbb{E}[E(S,T)] = \frac{d\,|S|\,|T|}{|G|},$$

since each vertex in S has d neighbors, and on average a fraction of $|T|/|G|$ of these neighbors will be in T. The expander mixing lemma gives a bound for how far $E(S,T)$ can be from its expectation in terms of the sizes of S and T, and the spectrum of the adjacency matrix.

Lemma 6.3.1 (Expander Mixing Lemma)**.** *Suppose* G *is a connected* k*-regular graph whose nontrivial adjacency eigenvalues are bounded in modulus by* λ*. If* S *and* T *are subsets of the vertex set of* G*, then*

$$\left| E(S,T) - \frac{d\,|S|\,|T|}{|G|} \right| \leqslant \lambda \sqrt{|S|\,|T| \left(1 - \frac{|S|}{|G|}\right)\left(1 - \frac{|T|}{|G|}\right)}.$$

Proof. Let $\mathbb{1}_S$ and $\mathbb{1}_T$ be the indicator vectors for S and T. Note that $E(S,T) = \mathbb{1}_S^T A \mathbb{1}_T$. The constant vector $\mathbb{1}$ is an eigenvector of A with eigenvalue k, and we expand the indicator vectors in terms of this eigenspace and its complement.

$$\mathbb{1}_S = \frac{|S|}{n}\mathbb{1} + \mathbb{1}_S^{\perp}; \quad \mathbb{1}_T = \frac{|T|}{n}\mathbb{1} + \mathbb{1}_T^{\perp}.$$

We then have

$$\begin{aligned} E(S,T) = \mathbb{1}_S^T A \mathbb{1}_T &= \left(\frac{|S|}{n}\mathbb{1} + \mathbb{1}_S^{\perp}\right)^T A \left(\frac{|T|}{n}\mathbb{1} + \mathbb{1}_T^{\perp}\right) \\ &= \frac{|S|\,|T|}{n^2}\mathbb{1}^T A \mathbb{1} + \frac{|T|}{n}(\mathbb{1}_S^{\perp})^T A \mathbb{1} + \frac{|S|}{n}\mathbb{1} A \mathbb{1}_T^{\perp} + (\mathbb{1}_S^{\perp})^T A \mathbb{1}_T^{\perp}. \end{aligned}$$

Since $\mathbb{1}$ is an eigenvector of A and $\mathbb{1} \perp \mathbb{1}_\bullet^{\perp}$, the middle two terms vanish. Further, $\mathbb{1}^T A \mathbb{1} = kn$, so we have $E(S,T) = \frac{k|S||T|}{n} + (\mathbb{1}_S^{\perp})^T A \mathbb{1}_T^{\perp}$. Taking absolute values, we have

$$\left| E(S,T) - \frac{k\,|S|\,|T|}{n} \right| = \left| (\mathbb{1}_S^{\perp})^T A \mathbb{1}_T^{\perp} \right| \leqslant \|A\mathbb{1}_T^{\perp}\| \|\mathbb{1}_S^{\perp}\| \leqslant \lambda \|\mathbb{1}_T^{\perp}\| \|\mathbb{1}_S^{\perp}\|.$$

It only remains to compute the norms of the complement vectors. We have $\|\mathbb{1}_S^{\perp}\|^2 + \frac{|S|^2}{n^2}\|\mathbb{1}\|^2 = \|\mathbb{1}_S\|^2$, giving $\|\mathbb{1}_S^{\perp}\| = \sqrt{|S| - \frac{|S|^2}{n}} = \sqrt{|S|\left(1 - \frac{|S|}{n}\right)}$. Applying the same treatment to $\mathbb{1}_T^{\perp}$ gives us our final bound. □

Note that there is no requirement that the graph be unweighted—regularity is all that is needed. In this case, the edges between S and T are counted in a weighted fashion. This suggests that it should be possible to generalize to matrix-valued weights as well. To do this, we need a matrix-valued analogue for $E(S,T)$. A reasonable choice is $E(S,T) = \sum_{u \in S, v \in T} A_{uv}$, where A is the block adjacency matrix of a matrix-weighted graph on n vertices with d-dimensional stalks. We can extract this information from the adjacency matrix in the same way as with the adjacency matrix of a standard graph. Let I_S be the $nd \times d$ matrix with a copy of the identity in each block corresponding to a vertex in S and zeros everywhere else, and similarly for I_T. Then $E(S,T) = I_S^T A I_T$, and we can use the same tricks

as for the standard expander mixing lemma. We decompose $I_S = \frac{|S|}{n} I_G + I_S^\perp$ and $I_T = \frac{|T|}{n} I_G + I_T^\perp$, where I_G is the $nd \times d$ matrix with a copy of the identity in every $d \times d$ block. Note that this is an orthogonal decomposition: every column of I_G is perpendicular to every column of $I_S^\perp$. Further, the columns of I_G and $I_S^\perp$ are mutually orthogonal, as can be seen by considering their supports. This allows us to perform the following decomposition:

$$\begin{aligned} E(S,T) = I_S^T A I_T &= \left(\frac{|S|}{n} I_G + I_S^\perp\right)^T A \left(\frac{|T|}{n} I_G + I_T^\perp\right) \\ &= \frac{|S|\,|T|}{n^2} I_G^T A I_G + \frac{|S|}{n} I_G A I_T^\perp + (I_S^\perp)^T A \frac{|T|}{n} I_G + (I_S^\perp)^T A I_T^\perp. \end{aligned}$$

Note now that every column of I_G is an eigenvector of A with eigenvalue k; since these are orthogonal to every column of $I_S^\perp$ and $I_T^\perp$, the two middle terms vanish. This fact also allows us to simplify the first term, giving

$$E(S,T) - \frac{k\,|S|\,|T|}{n} I = (I_S^\perp)^T A I_T^\perp.$$

At this point we can extract two bounds. One comes from taking the trace of this equation and taking absolute values:

$$\left|\operatorname{tr}(E(S,T)) - \frac{dk\,|S|\,|T|}{n}\right| = \left|\operatorname{tr}((I_S^\perp)^T A I_T^\perp)\right| \leqslant \sum_{i=1}^{d} |\lambda_{d+i}|\,\|(I_S^\perp)_i\|\|(I_T^\perp)_i\|,$$

where, e. g., $(I_S^\perp)_i$ is the ith column of $I_S^\perp$. These norms are straightforward to compute. For instance, $\|(I_S^\perp)_i\|^2 + \|\frac{|S|}{n}(I_G)_i\|^2 = \|(I_S)_i\|^2$, so

$$\|(I_S^\perp)_i\| = \sqrt{|S| - \frac{|S|^2}{n^2} n} = \sqrt{|S|\,(1 - \frac{|S|}{n})}.$$

Putting this all together, we have

$$\left|\operatorname{tr}(E(S,T)) - \frac{dk\,|S|\,|T|}{n}\right| \leqslant \sum_{i=1}^{d} |\lambda_{d+i}| \sqrt{|S|\,|T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)}.$$

Alternately, instead of bounding the trace, we can produce a bound in the semidefinite order. Let $\left|(I_S^\perp)^T A I_T^\perp\right|$ be the positive semidefinite matrix obtained

by replacing all eigenvalues of $(I_S^\perp)^T A I_T^\perp$ with their absolute values. In particular, then, we have

$$-\left|(I_S^\perp)^T A I_T^\perp\right| \preceq E(S,T) - \frac{k|S||T|}{n} I \preceq \left|(I_S^\perp)^T A I_T^\perp\right|.$$

Meanwhile, $\left|(I_S^\perp)^T A I_T^\perp\right| \preceq \|(I_S^\perp)^T A|_{V_1^\perp} I_T^\perp\| I$, where $A|_{V_1^\perp}$ is the adjacency operator restricted to the orthogonal complement of the eigenspace of constant cochains, and

$$\|(I_S^\perp)^T A|_{V_1^\perp} I_T^\perp\| \leqslant \|I_S^\perp\| \|A|_{V_1^\perp}\| \|I_T^\perp\| = |\lambda_{d+1}| \sqrt{|S||T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)}.$$

In summary, we have

Lemma 6.3.2 (Expander mixing lemma for regular matrix-weighted graphs)**.** *Let $\mathcal{F}$ be a kI-regular matrix-weighted graph on an n-vertex graph G, whose vertex stalks are dimension d. If S and T are subsets of the vertices of G, then the following hold:*

$$\left|\operatorname{tr}(E(S,T)) - \frac{dk|S||T|}{n}\right| \leqslant \sum_{i=1}^{d} |\lambda_{d+i}| \sqrt{|S||T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)} \tag{6.1}$$

$$E(S,T) - \frac{k|S||T|}{n} I \preceq |\lambda_{d+1}| \sqrt{|S||T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)} I, \tag{6.2}$$

$$E(S,T) - \frac{k|S||T|}{n} I \succeq -|\lambda_{d+1}| \sqrt{|S||T| \left(1 - \frac{|S|}{n}\right)\left(1 - \frac{|T|}{n}\right)} I \tag{6.3}$$

where $k = \lambda_1 = \cdots = \lambda_d \geqslant |\lambda_{d+1}| \geqslant \cdots$ are the eigenvalues of the adjacency matrix of $\mathcal{F}$ ordered by decreasing absolute value.

We can interpret this as saying that for a kI-regular matrix-weighted graph with good expansion constant, $E(S,T)$ is close to what its expectation would be for the constant sheaf $\underline{\mathbb{R}^d}$ on a random k-regular graph.

6.3.1 *Irregular graphs*

A similar result holds for irregular matrix-weighted graphs. For this we use the normalized adjacency matrix, which is given by the off-diagonal blocks of the

normalized Laplacian matrix of the sheaf. Equivalently, this is the matrix $\tilde{A} = D^{-1/2}AD^{-1/2}$. Define the $nd \times d$ matrix ψ whose blocks consist of the diagonal blocks of $D^{1/2}$. The columns of ψ are all eigenvectors of $\tilde{A}$ with eigenvalue 1, and

$$E(S,T) = I_S^T A I_T = \psi_S^T D^{-1/2} A D^{-1/2} \psi_T = \psi_S^T \tilde{A} \psi_T,$$

where ψ_S is the matrix equal to ψ in blocks corresponding to vertices in S and zero otherwise.

In irregular graphs, it is more convenient to measure the size of a set of vertices by its volume: the sum of the degrees of the vertices in the set. Similarly, for irregular matrix-weighted graphs sheaves, we define $\operatorname{vol}(S)$ to be the sum of the degree matrices of the vertices in S. Then $\operatorname{vol}(S) = I_S^T D I_S = \psi_S^T \psi_S = \psi_S^T \psi$. We now decompose $\psi_S = \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S) + \psi_S^\perp$. The two terms in this decomposition are orthogonal in the sense that

$$\begin{aligned}(\psi_S^\perp)^T \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S) &= (\psi_S - \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S))^T \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S) \\ &= \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(S) - \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(G) \operatorname{vol}(G)^{-1} \operatorname{vol}(S) = 0.\end{aligned}$$

However, the individual columns of the two matrices are not orthogonal. This means we will not get quite as tight a bound as in the regular case. However, we do have

$$\begin{aligned}E(S,T) &= (\psi_S)^T \tilde{A} \psi_T \\ &= (\psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S) + \psi_S^\perp)^T \tilde{A} (\psi \operatorname{vol}(G)^{-1} \operatorname{vol}(T) + \psi_T^\perp) \\ &= \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \psi \tilde{A} \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(T) + (\psi_S^\perp)^T \tilde{A} \psi_T^\perp \\ &= \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(T) + (\psi_S^\perp)^T \tilde{A} \psi_T^\perp.\end{aligned}$$

Therefore, we have

$$E(S,T) - \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(T) = (\psi_S^\perp)^T \tilde{A} \psi_T^\perp.$$

Taking the trace gives

$$\begin{aligned}\left|\operatorname{tr}(E(S,T) - \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(T)\right| &\leqslant \left|\operatorname{tr}((\psi_S^\perp)^T \tilde{A} \psi_T^\perp)\right| \\ &\leqslant \|\psi_S^\perp\|_F \|\tilde{A} \psi_T^\perp\|_F \leqslant |\lambda_{d+1}| \, \|\psi_S^\perp\|_F \|\psi_T^\perp\|_F.\end{aligned}$$

The norms in this formula are, e. g.,

$$\begin{aligned}\|\psi_S^{\perp}\|_F &= \operatorname{tr}\left[(\psi_S - \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S))^T(\psi_S - \psi \operatorname{vol}(G)^{-1} \operatorname{vol}(S))\right] \\ &= \operatorname{tr}[\operatorname{vol}(S) + \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(S) \\ &- \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(S) - \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(S)] \\ &= \operatorname{tr}\left[\operatorname{vol}(S) - \operatorname{vol}(S) \operatorname{vol}(G)^{-1} \operatorname{vol}(S)\right]\end{aligned}$$

Combining this all gives

Lemma 6.3.3 (Expander Mixing Lemma for irregular matrix-weighted graphs)**.** *Let $\mathcal{F}$ be a matrix-weighted graph with underlying unweighted graph G, with vertex stalks dimension d. If S and T are subsets of the vertices of G, then*

$$|\operatorname{tr}[E(S,T) - V(S,T)]| \leqslant |\lambda_{d+1}| \sqrt{\operatorname{tr}(\operatorname{vol}(S) - V(S,S)) \operatorname{tr}(\operatorname{vol}(T) - V(T,T))},$$

where $V(A,B) = \operatorname{vol}(A) \operatorname{vol}(G)^{-1} \operatorname{vol}(B)$ and where $1 = \lambda_1 = \ldots = \lambda_d \geqslant |\lambda_{d+1}| \geqslant \ldots$ are the eigenvalues of the normalized adjacency matrix of $\mathcal{F}$ ordered by decreasing absolute value.

6.4 THE CHEEGER INEQUALITY

There are two spectral results typically associated with notions of expansion in a graph. One is the expander mixing lemma already discussed, which can be seen as relating the two-sided spectral expansion of a graph with its edge expansion. The other is the Cheeger inequality, which relates the one-sided spectral expansion with the edge expansion.

Our definition of $E(S,T)$ for two subsets of vertices in a matrix-weighted graph suggests a way to generalize the Cheeger inequality. We define a Cheeger constant for a subset S of the vertices of a kI-regular matrix-weighted graph as

$$h(S) = \frac{\operatorname{tr}\left[E(S,\overline{S})\right]}{\min(|S|, |\overline{S}|)}.$$

We can compute $E(S,\overline{S})$ using the Laplacian matrix:

$$E(S,\overline{S}) = I_S^T L I_S = \left(I_S - \frac{|S|}{n} I_G\right)^T L \left(I_S - \frac{|S|}{n} I_G\right),$$

while $(I_S - \frac{|S|}{n} I_G)^T (I_S - \frac{|S|}{n} I_G) = (|S| + \frac{|S|^2}{n} - 2\frac{|S|^2}{n})I = \frac{|S||\overline{S}|}{n} I$.

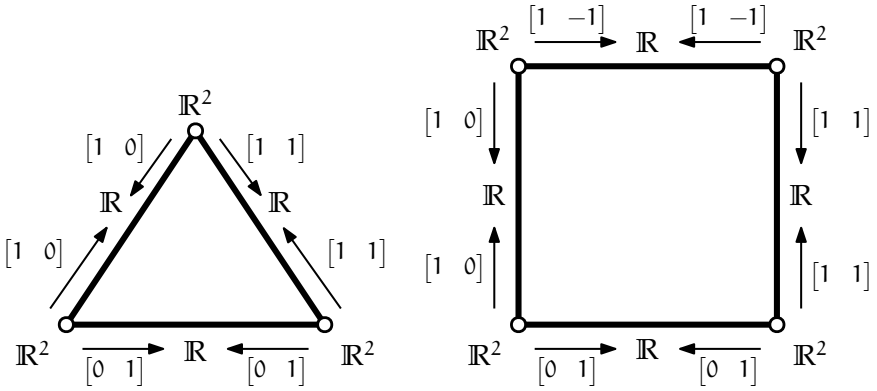


Figure 6.1: Some counterexamples to the right hand side of a sheaf Cheeger inequality.

Applying a generalization of the Courant-Fischer-Weyl theorem, we have

$$\sum_{i=1}^{d} \lambda_{d+i} \leqslant \frac{\operatorname{tr}(I_S^T L I_S)}{\operatorname{tr}((I_S - \frac{|S|}{n} I_G)^T (I_S - \frac{|S|}{n} I_G))} = \frac{\operatorname{tr} E(S, \overline{S})}{d\frac{|S||\overline{S}|}{n}} \leqslant 2\frac{\operatorname{tr} E(S, \overline{S})}{d \min(|S|, |\overline{S}|)},$$

where we order the eigenvalues λ_i of L in increasing order. This is the "easy" side of a Cheeger-type inequality:

$$\frac{1}{2}\sum_{i=1}^{d} \lambda_{d+i} \leqslant \frac{h_G}{d}.$$

To get an upper bound on this Cheeger constant, we need a way to round an eigenvector (or block of eigenvectors) to an indicator matrix on vertices of the graph.

Unfortunately, this is not possible. To be more precise: there exists a matrix-weighted graph $\mathcal{F}$ with a nonconstant section x such that x cannot be rounded to a vector of the form $v \otimes y$ for some v in the vertex stalk. Consider the sheaf illustrated in Figure 6.1.

Note that there is a nonconstant section whose values on vertex stalks are, counterclockwise from the lower right,

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \text{ and } \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Suppose there were a vector of the form $v \otimes y$, for some $v \in \mathbb{R}^2$ and $y \in \mathbb{R}^3$, with $y \perp \mathbb{1}$, such that $L_{\mathcal{F}}(v \otimes y) = 0$. Consider the graph G_v whose edge weights

are $v^T A_{ij} v$. Note then that $(v \otimes y)^T L_{\mathcal{F}}(v \otimes y) = y^T L_{G_v} y$, where L_{G_v} is the graph Laplacian of G_v. Then there exists some y such that $L_{\mathcal{F}}(v \otimes y) = 0$ only if G_v is not connected. This requires that there be some cut set of the underlying graph such that v lies in the kernel of every restriction map corresponding to an edge in the cut. We can see by inspection that no such v exists for the sheaf under consideration.

A similar argument applies to the second sheaf in Figure 6.1. This sheaf has a two-dimensional space of non-constant sections, but no sections which are constant and supported on a subgraph. An indicator matrix I_S for a proper subset of vertices of the graph such that $\mathrm{tr}(I_S^T L_{\mathcal{F}} I_S) = 0$ exists if and only if the graph with edge weights $\mathrm{tr}(A_{ij})$ is disconnected. But this only happens if there is already a cutset of edges whose edge weights are $A_{ij} = 0$.

The observant reader will note that these counterexamples are quite similar to the counterexample given to the converse to Proposition 5.2.1 on a necessary but not sufficient condition for a matrix-weighted graph to be an approximation to the constant sheaf. In a sense, the left-hand Cheeger inequality is a spectral version of this condition, and the failure of the right-hand inequality to hold results from the existence of sheaves which satisfy the conclusion of Proposition 5.2.1 but fail to be approximations to the constant sheaf.

The conclusion we must draw is that for matrix-weighted graphs, combinatorial measures of expansion will tend to be weaker than spectral measures of expansion. We cannot control the spectrum of a matrix-weighted graph by looking only at edge cuts.

7

DYNAMICS

7.1 THE HEAT EQUATION

The Laplacian on a domain in $\mathbb{R}^n$ or on a Riemannian manifold has many associated partial differential equations. The simplest, the Laplace equation $\Delta f = 0$, inspired our study of harmonic extension in Section 3.1. Adding a time dimension brings us to the heat equation $\frac{df}{dt} = -\Delta f$. These also have immediate analogues for sheaf Laplacians. The homological properties of the sheaf Laplacians means trajectories have a nice interpretation.

Proposition 7.1.1. *Let $x(t)$ be a solution to the sheaf heat equation $\frac{dx}{dt} = -\Delta^k_{\mathcal{F}} x$ with initial condition $x(0) = x_0$. Then $\lim_{t\to\infty} x(t)$ is equal to the harmonic k-cochain nearest to x_0 in the ℓ^2 sense.*

Proof. We may write $x(t) = e^{-t\Delta^k_{\mathcal{F}}} x_0$. The sheaf Laplacian $\Delta^k_{\mathcal{F}}$ is, by construction, positive semidefinite, and we can diagonalize it into $\Delta^k_{\mathcal{F}} = U^* D U$ for U unitary and D diagonal and nonnegative. Then $e^{-t\Delta^k_{\mathcal{F}}} = U^* e^{-tD} U$. As $t \to \infty$, e^{-tD} converges to a diagonal matrix with zeros corresponding to every positive entry of D and a 1 corresponding to every zero entry of D. Thus, $e^{-t\Delta^k_{\mathcal{F}}}$ converges to the orthogonal projection onto the kernel of $\Delta^k_{\mathcal{F}}$, so $x(t)$ converges to the orthogonal projection of x_0 onto $\mathcal{H}^k(X;\mathcal{F})$. □

Naturally, similar results hold for differential equations using only the up- or down-Laplacians. Since the sheaf Laplacian may be computed locally—to compute $(\Delta^k x)_\sigma$ one only needs information about x at cells which are adjacent to it in the cell complex—this offers a distributed way to find representatives of sheaf cohomology classes. This property of the heat equation will be the foundation of many applications in Part II. The rate of convergence to a global section is governed by the smallest nonzero eigenvalue of Δ^k. The larger this eigenvalue, the faster the convergence.

7.1.1 *Boundary conditions*

We may also solve the heat equation with boundary conditions: choose a subcomplex $B \subseteq X$ and fix a k-cochain on B, then use the Laplacian to determine derivatives

for cells not in B. Perhaps unsurprisingly, trajectories of the heat equation converge to the harmonic extension of their boundary conditions. If $U = X \setminus B$, the relevant differential equation may be written

$$\frac{d}{dt} x_U = -\Delta^k[U,U]x_U - \Delta^k[U,B]x_B.$$

The stationary points of this differential equation are precisely where $(\Delta^k x)_U = 0$. Linearizing about a stationary point turns the equation into

$$\frac{d}{dt} x_U = -\Delta^k[U,U]x_U,$$

which is stable, because $\Delta^k[U,U]$ is a principal submatrix of a positive semidefinite matrix and hence is positive semidefinite. If the harmonic extension of x_B is not unique, x will converge to the harmonic extension of x_B nearest the initial condition, since $\ker \Delta^k[U,U]$ is preserved by the flow. That is, $\dot{x_U} \perp \ker \Delta^k[U,U]$, since $\operatorname{im} \Delta^k[U,B] \subseteq \operatorname{im} \Delta^k[U,U] \perp \ker \Delta^k[U,U]$. Since the portion of the initial condition lying in $\ker \Delta^k[U,U]$ is undisturbed, the limit is the closest k-cochain possible satisfying the harmonic extension.

7.1.2 *Input-Output dynamics*

Suppose we have a linear system

$$\dot{x} = -\Delta^k_{\mathcal{F}} x + i_B u \tag{7.1}$$

$$y = \pi_C x \tag{7.2}$$

where u is an input variable and y an output variable, and $i_B : C^k(B;\mathcal{F}) \to C^k(X;\mathcal{F})$ is the inclusion map of a subcomplex, and $\pi_C : C^k(X;\mathcal{F}) \to C^k(C;\mathcal{F})$ is the projection map onto a subcomplex. This is a dynamical system on X with inputs given at cells in B and outputs observed at cells in C. What properties does this control system have? One such desirable property is controllability: given a target state x_T, does there exist an input function $u(t)$ that steers the state to x_T in a finite amount of time? A slightly looser condition is *stabilizability*, which ensures controllability in the limit. A system is stabilizable if any subspace of the state space which is not asymptotically stable to the origin is controllable. For the heat equation, most of the state space is asymptotically stable, and we only need to worry about the kernel of Δ^k.

Proposition 7.1.2. *The system* (7.1) *is stabilizable if* $H^k(X,B;\mathcal{F}) = 0$.

Proof. Stabilizability is equivalent to the condition that the matrix

$$\begin{bmatrix}(-\Delta_{\mathcal{F}}^{k} - \lambda I) & i_B\end{bmatrix}$$

be full rank for all λ with nonnegative real part. The eigenvalues of $-\Delta_{\mathcal{F}}^{k}$ are real and nonpositive, so this matrix is already full rank for all necessary values of λ except possibly 0. Suppose now that $H^k(X, B; \mathcal{F}) = 0$. This implies that i_B induces a surjective map $H^k(B; \mathcal{F}) \to H^k(X; \mathcal{F})$, which means that for any $y \in \ker \Delta_{\mathcal{F}}^{k} \simeq H^k(X; \mathcal{F})$, there exists some $x \in C^k(B; \mathcal{F})$ such that $\langle x, y\rangle \neq 0$, and hence $\operatorname{im} i_B$ and $\operatorname{im} \Delta_{\mathcal{F}}^{k}$ span $C^k(X; \mathcal{F})$, so that the required matrix is full rank. □

Every statement about linear control systems has a dual statement about the dual system. Since the Laplacian control system is essentially self-dual (simply swap the roles of B and C), we have the corresponding result

Proposition 7.1.3. *The system* (7.1) *is detectable if* $H^k(X, C; \mathcal{F}) = 0$.

Detectability is a generalization of observability—it requires that we be able to determine the state of the system up to an element of the stable subspace. In this case, it means that we can deduce the nearest harmonic k-cochain to the state. Combining detectability and stabilizability means we can design an observer-controller pair that will steer the system asymptotically to any desired harmonic k-cochain.

7.2 DYNAMICS OF SHEAVES

The algebraic structure of a cellular sheaf, described by coboundary matrices, offers a new nu opportunity for dynamics: rather than modify a cochain in a way depending on the sheaf structure, we can modify the sheaf itself based on some known cochain. We will begin with the simplest case: a sheaf $\mathcal{F}$ on a graph G. Consider the dynamics on restriction maps of $\mathcal{F}$ given by

$$\frac{d}{dt}\mathcal{F}_{v \trianglelefteq e} = -(\delta_{\mathcal{F}} x)_e x_v^*. \tag{7.3}$$

Proposition 7.2.1. *The trajectories of* 7.3 *initialized at a sheaf* $\mathcal{F}_0$ *on* G *converge to a sheaf* $\mathcal{F}'$*, which is the minimizer of* $d(\mathcal{F}_0, \mathcal{F}')$ *subject to the constraint that* $x \in H^0(G; \mathcal{F})$*, where the distance between sheaves is measured by*

$$d(\mathcal{F}, \mathcal{F}') = \|\delta_{\mathcal{F}} - \delta_{\mathcal{F}'}\|_F^2 = \sum_{\sigma \trianglelefteq \tau} \|\mathcal{F}_{\sigma \trianglelefteq \tau} - \mathcal{F}'_{\sigma \trianglelefteq \tau}\|_F^2.$$

Proof. The stationary points of this dynamical system are precisely those where $x \in \ker \delta^0_{\mathcal{F}}$, that is, those defining a sheaf for which x is a global section.. For a given edge e, the derivative of δ^0_e, the block row of $\delta^0_{\mathcal{F}}$ corresponding to e, is equal to $-\delta_e x_{\partial e} x^*_{\partial e}$, where $x_{\partial e} = \sum_{v \trianglelefteq_1 e} x_e \in C^0(G;\mathcal{F})$. The matrix $x_{\partial e} x^*_{\partial e}$ is positive semidefinite, and acts on the space of matrices δ_e with the correct sparsity pattern by $\delta_e \mapsto \delta_e x_{\partial e} x^*_{\partial e}$. The kernel of this operator is precisely the set of δ_e for which $\delta_e x = 0$; equivalently, each row of δ_e must be in the kernel of $x_{\partial e} x^*_{\partial e}$. Since the dynamics act on each row of δ_e separately, we apply the same argument as for Proposition 7.1.1 to see that the limit is the projection of each row of δ_e onto the kernel of $x_{\partial e} x^*_{\partial e}$, and hence the projection of δ_e onto this kernel in the Frobenius norm. Since there is no interaction between edges and the squared Frobenius norm is additive, this is the minimizer of the total distance. □

The sheaf dynamics equation (7.3) is a sort of dual to the heat equation. It is a linear diffusion acting independently on the restriction maps corresponding to each edge. We can construct similar dynamics with multiple 0-cochains x_i, with similar behavior where each x_i is a global section of the limit. Of course, it may be the case that the limiting sheaf is in fact the sheaf with all restriction maps zero.

Dually to (7.3), we might consider some $y \in C^1(G;\mathcal{F})$ and the dynamics

$$\frac{d}{dt}\mathcal{F}_{v \trianglelefteq e} = -y_e y_e^* \mathcal{F}_{v \trianglelefteq e}.$$

In the limit, y will be in $\mathcal{H}^1(G;\mathcal{F})$, but we have no guarantees about optimality. All that this differential equation does is modify $\mathcal{F}$ so that y_e becomes orthogonal to $\operatorname{im} \mathcal{F}_{v \trianglelefteq e}$ for all e. This guarantees that y is orthogonal to $\operatorname{im} \delta^0$, but this is clearly not in general the nearest sheaf satisfying this condition.

Things get a bit trickier when we consider higher-dimensional complexes. While the requirement that restriction maps commute is trivial when the base space is a graph, this constraint becomes trickier to manage as the dimensionality increases. If we restrict ourselves to modifying a single layer of restriction maps, however, the commutativity constraint remains a linear condition. Indeed, the condition that $\delta^k \delta^{k-1} = 0$ for all k is equivalent to the commutativity of the diagram a cellular sheaf defines in **Vect**. To see this, note that

$$(\delta^k \delta^{k-1} x_\sigma)_\tau = \sum_{\sigma \trianglelefteq_1 \chi \trianglelefteq_1 \tau} [\sigma : \chi][\chi : \tau] \mathcal{F}_{\chi \trianglelefteq \tau} \mathcal{F}_{\sigma \trianglelefteq \chi} x_\sigma.$$

There are precisely two cells χ lying between σ and τ, and the conditions for the incidence relation ensure that their corresponding terms have coefficients of opposite signs, so this equation is equivalent to $\mathcal{F}_{\chi \trianglelefteq \tau} \mathcal{F}_{\sigma \trianglelefteq \chi} x_\sigma = \mathcal{F}_{\chi' \trianglelefteq \tau} \mathcal{F}_{\sigma \trianglelefteq \chi'} x_\sigma$,

which, if it holds for all x_σ, is precisely the requirement for the restriction map $\mathcal{F}_{\sigma \trianglelefteq \tau}$ to be well defined. As a result, if we wish to modify the restriction maps collated in δ^k while preserving the fact that these define a cellular sheaf, we only need to preserve the equations $\delta^k \delta^{k-1} = 0$ and $\delta^{k+1}\delta^k = 0$.

Suppose X is a $(k+1)$-dimensional complex, and consider the dynamics

$$\frac{d}{dt}\delta^k_\sigma = -\delta^k_\sigma(\Delta^k_-[\partial\sigma, \partial\sigma] + x_{\partial\sigma}x^*_{\partial\sigma}), \tag{7.4}$$

which for $k = 0$ reduces to (7.3).

Proposition 7.2.2. *Suppose $x \in \ker \Delta^k_-$. The trajectories of (7.4) initialized at a sheaf $\mathcal{F}_0$ converge to the nearest sheaf to $\mathcal{F}_0$ such that $x \in \ker \Delta^k$.*

Proof. First note that the stationary points of this dynamical system are precisely those where $\delta^k\delta^{k-1} = 0$ and $\delta^k x = 0$. Naturally $\delta^k x = 0$ if and only if $\delta^k_\sigma x_{\partial\sigma} x^*_{\partial\sigma} = 0$ for all σ of dimension $k+1$. Similarly, $\delta^k\delta^{k-1} = 0$ if and only if $\delta^k_\sigma \delta^{k-1}_{\partial\sigma}(\delta^{k-1}_{\partial\sigma})^* = \delta^k_\sigma \Delta^k_-[\partial\sigma, \partial\sigma] = 0$ for all σ. If $x \in \ker \Delta^k_-$, then $x_{\partial\sigma} \in \ker \Delta^k_-[\partial\sigma, \partial\sigma]$ for all σ. (This fact relies on the structure of the cell complex and is not true if we replace Δ^k_- with Δ^k_+.) But then this means that $\ker(\Delta^k_-[\partial\sigma, \partial\sigma] + x_{\partial\sigma}x^*_{\partial\sigma}) = \ker(\Delta^k_-[\partial\sigma, \partial\sigma]) \cap \ker(x_{\partial\sigma}x^*_{\partial\sigma})$, so that δ^k is an equilibrium of (7.4) if and only if $\delta^k\delta^{k-1} = 0$ and $\delta^k x = 0$.

By our previous considerations, then, trajectories converge to the orthogonal projection onto the kernel of this operator; this is the intersection of the space of δ^k defining commuting restriction maps given δ^{k-1} and the space of δ^k with $\delta^k x = 0$. Thus, the limit is the nearest sheaf to $\mathcal{F}_0$ with $x \in \ker \Delta^k$. □

Of course, it is possible to combine the two processes, for a diffusion process acting jointly on the structure maps and the data:

$$\begin{aligned} \frac{d}{dt}x &= -\alpha(\delta^k)^*\delta^k x \\ \frac{d}{dt}\delta^k_\sigma &= -\beta\delta^k_\sigma(\Delta^k_-[\partial\sigma, \partial\sigma] + x_{\partial\sigma}x^*_{\partial\sigma}) \end{aligned} \tag{7.5}$$

These systems are most immediately interesting when applied to 0-cochains on a sheaf on a graph G. One process tries to find the nearest global section to x_0, while the other tries to find the nearest sheaf for which x is a global section. When $\alpha = \beta = 1$, this is gradient descent on the function

$$V(x, \delta) = \frac{1}{2}\langle \delta x, \delta x\rangle$$

with respect to both x and δ, where the sparsity pattern of δ is constrained to respect the underlying graph. For arbitrary $\alpha, \beta \geqslant 0$, V is still a Lyapunov function:

$$\begin{aligned}\dot{V}(x,\delta) &= x^*\delta^*\delta\dot{x} + x^*\delta^*\dot{\delta}x \\ &= -\alpha x^*\delta^*\delta\delta^*\delta x - \beta x^*\delta^* P_\delta(\delta x x^*)x \\ &= -\alpha x^*(\delta^*\delta)^2 x - \beta \sum_e x^*_{\partial e}\delta^*_e\delta_e x_{\partial e} x^*_{\partial e} x_{\partial e}.\end{aligned}$$

Here, P_δ is the projection from $\mathrm{Hom}(C^0, C^1)$ to the space of matrices with the correct sparsity pattern to be a coboundary map for a sheaf on G. The first term is nonpositive, and is negative whenever $\delta x \neq 0$, and the second term is negative whenever $\delta_e x_{\partial e} \neq 0$ for some e. Thus $\dot{V}$ is negative except on the set of x and δ where $\delta x = 0$, where it vanishes. Thus, trajectories of the combined dynamics converge to pairs consisting of a sheaf and a section.

The set of stationary points forms a degree-2 algebraic variety; since this is a connected set, there is no hope of global exponential stability of any given equilibrium. However, nonsingular points in this variety are Lyapunov stable. This is straightforward: the nonsingular points form a manifold, which is equal to the center manifold of the dynamical system at these points.

For most graphs and choices of stalk dimensions, "most" points of the variety of stationary points have $\ker \delta = 0$, since there are many more full-rank choices for δ than rank-deficient ones. However, there are many initial conditions whose trajectories converge to a δ with nontrivial kernel, i. e., a sheaf with nontrivial sections.

Proposition 7.2.3. *If one of the diagonal blocks of $\alpha\delta_0^*\delta_0 - \beta x_0 x_0^\mathsf{T}$ fails to be positive semidefinite, the trajectory of (7.5) converges to a point (x, δ) with $x \neq 0$.*

Proof. Write $M = \alpha\delta^*\delta - \beta x x^*$ and consider

$$\begin{aligned}\frac{d}{dt}M &= \frac{d}{dt}\alpha\delta^*\delta - \beta x x^* \\ &= \alpha(-\beta P_\delta(\delta x x^*)^*\delta - \beta\delta^* P_\delta(\delta x x^*)) - \beta(-\alpha x x^*\delta^*\delta - \alpha\delta^*\delta x x^*) \\ &= -\alpha\beta[(\delta^* P_\delta(\delta x x^*))^* + \delta^* P_\delta(\delta x x^*)] + \alpha\beta[(\delta^*\delta x x^*)^* + \delta^*\delta x x^*].\end{aligned}$$

The diagonal block of $\delta^* P_\delta(\delta x x^*)$ corresponding to a vertex v is equal to $(\delta^*\delta x x^*)[v, v]$, since the sparsity pattern of the relevant block row of δ^* is the same as the sparsity pattern imposed by the projection P_δ. Thus, the diagonal blocks of the two terms in the derivative cancel, giving $\frac{d}{dt}\operatorname{diag}(M) = 0$. Hence, if $M[v, v]$ is indefinite at $t = 0$, it is indefinite for all t, which implies that xx^* cannot go to zero, since otherwise M would converge to $\delta^*\delta$, which has by definition positive semidefinite diagonal blocks. □

In particular, this ensures that for a given initial δ and x, we can always rescale the initial condition (or the weighting parameters α and β) so that the limiting sheaf has a nontrivial global section. We can also control some other quantities during the evolution of these combined dynamics.

Proposition 7.2.4. *The quantities* $\|\delta\|_F^2$, $\|x\|^2$, $\|\delta x\|^2$, *and* $\frac{\|\delta x\|^2}{\|x\|^2}$ *are nonincreasing under* (7.5).

Proof. Our Lyapunov function for these dynamics is $V = \frac{1}{2}\|\delta x\|^2$, so this quantity is clearly nonincreasing. We evaluate the other derivatives:

$$\begin{aligned}\frac{d}{dt}\|\delta\|_F^2 &= \frac{d}{dt}\operatorname{tr}(\delta^*\delta)\\ &= -\beta \operatorname{tr}(\delta^* P_\delta(\delta x x^*) + P_\delta(\delta x x^*)^*\delta)\\ &= -2\beta \sum_e (x_e^*\delta_e^*\delta_e x_e) x_e^* x_e \leqslant 0.\end{aligned}$$

$$\begin{aligned}\frac{d}{dt}\|x\|^2 &= \frac{d}{dt}x^*x\\ &= -\alpha(x^*\delta^*\delta x + (\delta^*\delta x)^* x)\\ &= -2\alpha x^*\delta^*\delta x \leqslant 0\end{aligned}$$

Combining these, we have

$$\frac{d}{dt}\frac{\|\delta x\|^2}{\|x\|^2} = \frac{\|x\|^2 \frac{d}{dt}\|\delta x\|^2 - \|\delta x\|^2 \frac{d}{dt}\|x\|^2}{\|x\|^4}.$$

This is bounded above by

$$\frac{-2\alpha\|x\|^2(\alpha x^*(\delta^*\delta)^2 x) + \|\delta x\|^2 2\alpha(x^*\delta^*\delta x)}{\|x\|^4} \leqslant 2\alpha\left[\left(\frac{x^*\delta^*\delta x}{\|x\|^2}\right)^2 - \frac{x^*(\delta^*\delta)^2 x}{\|x\|^2}\right].$$

To see that the value inside the brackets is always negative, take $\|x\| = 1$ and diagonalize $\delta^*\delta$ to get the equivalent inequality

$$\sum_i \lambda_i^2 x_i \geqslant \left(\sum_i \lambda_i x_i\right)^2,$$

for $\lambda_i, x_i \geqslant 0$ and $\sum_i x_i = 1$; this is simply Jensen's inequality for the convex function $\lambda \mapsto \lambda^2$. □

The fact that $\frac{\|\delta x\|^2}{\|x\|^2}$ is decreasing means that x becomes a section of $\mathcal{F}$ at least as quickly as it approaches zero.

Proposition 7.2.5. *Let (x_0, δ_0) be a point with $\delta_0 x_0 = 0$, such that for some neighborhood U of (x_0, δ_0), the set of stationary points of (7.5) intersected with U is a manifold. Then (7.5) is Lyapunov stable at (x_0, δ_0).*

Proof. We can identify the stationary points in U with the center manifold at (x_0, δ_0). The tangent space at (x_0, δ_0) to the space of stationary points is $\ker D(\delta x)$, since the space of stationary points is precisely those for which $\delta x = 0$. We have $D(\delta x) = \delta_0 x + \delta x_0)$, which characterizes $\ker D(\delta x)$. Meanwhile, the linearization of (7.5) at (x, δ) is

$$\begin{aligned}\dot{x} &= -\alpha(\delta_0^* \delta x_0 + \delta_0^* \delta_0 x)\\ \dot{\delta_e} &= -\beta\left[(\delta_e)^*(x_0)_{\partial e}(x_0)^*_{\partial e} + (\delta_0)^*_e x_{\partial e}(x_0)^*_{\partial e}\right]\end{aligned}$$

Thus $\dot{x} = 0$ if and only if $(x, \delta) \in \ker D(\delta x)$ and similarly for $\dot{\delta}$, and hence the stationary points form a center manifold at (x_0, δ_0). The center manifold theorem [Kel67] then says that stability at this point is equivalent to stability when restricted to the center manifold. But the dynamics on the center manifold are constant, and hence Lyapunov stable. □

One would expect Lyapunov stability to also hold at singular points of the variety of stationary points, but the center manifold theorem is not sufficient to show this.

7.3 NONLINEAR LAPLACIANS

The heat equation on a graph can be thought of in terms of a quadratic potential on each edge of the graph, pushing the values at the vertices toward the minimizer for that potential. That is, since Lx is the gradient of $\frac{1}{2}\|\delta x\|^2$, the heat equation is gradient descent with respect to this function. What happens when the potential function is nonquadratic, or perhaps even nonconvex? Consider a general potential function $V(x) = \frac{1}{2}\sum_e \phi_e((\delta x)_e)$, for some set of functions $\phi_e : \mathcal{F}(e) \to \mathbb{R}$. The differential of this function with respect to x is

$$dV(x)(y) = \sum_e (\nabla \phi_e)_{(\delta x)_e}(\delta y)_e.$$

Its gradient is then represented by $\nabla V_x = \delta^* \circ \nabla\phi \circ \delta_x$, where $\nabla\phi$ is the gradient of $\phi : C^1 \to \mathbb{R}$ given by $\phi(y) = \sum_e \phi_e(y_e)$. We can interpret ∇V as a sort of nonlinear degree-0 Laplacian, which we will denote $L^{\phi}_{\mathcal{F}}$. This nonlinear Laplacian

maintains its local computability, since it may be calculated by simply adding an extra edgewise computation between the components of the linear Laplacian.

If each ϕ_e is radially symmetric, with $\phi_e(y_e) = \psi_e(\|y_e\|^2)$ for some $\psi_e : [0,\infty) \to \mathbb{R}$, then $\nabla\phi_e(y_e) = \psi_e'(\|y_e\|^2)y_e$. In this situation the nonlinearity in $L_{\mathcal{F}}^{\Phi}$ amounts to a scaling factor of $\psi_e'(\|(\delta x)_e\|^2)$ on each edge. If $\psi_e'(z)$ is zero for $z \geqslant D$, the nonlinear Laplacian ignores sufficiently large discrepancies over edges. The corresponding gradient descent is a thresholded consensus dynamics on cochains of $\mathcal{F}$.

7.3.1 *Threshold dynamics*

Proposition 7.3.1. *Suppose we have $\phi_e(y_e) = \psi_e(\|y_e\|^2)$ for some $\psi_e : [0,\infty) \to \mathbb{R}$, with $\psi_e'(z) = 0$ for $z \geqslant D_e$ and $\psi_e'(z) > 0$ for $0 < z < D_e$. Then, for any $x \in C^0(G;\mathcal{F})$, $L_{\mathcal{F}}^{\Phi}(x) = 0$ if and only if either $x \in H^0(G;\mathcal{F})$ or, for every edge e with with $\delta_e x \neq 0$, $\|\delta_e x\|^2 \geqslant D_e$.*

Proof. If $x \in H^0(G;\mathcal{F})$, then $\delta x = 0$, and hence $\nabla\phi(\delta x) = 0$, since $\nabla\phi(0) = 0$. More generally, if $\|\delta_e x\|^2 \geqslant D_e$, then $\nabla\phi_e(\delta_e x) = \psi_e'(\|\delta_e x\|^2)\delta_e x = 0$.

On the other hand, if $L_{\mathcal{F}}^{\Phi}(x) = 0$, we have $\nabla\phi(\delta x) \in \ker \delta^*$, or equivalently, $\nabla\phi(\delta x) \perp \operatorname{im}\delta$. In particular, we must have $\langle \nabla\phi(\delta x), \delta x\rangle = 0$. Setting $y = \delta x$, we have

$$\langle \nabla\phi(y), y\rangle = \sum_e \langle \nabla\phi_e(y_e), y_e\rangle = \sum_e \langle \psi_e'(\|y_e\|^2)y_e, y_e\rangle = \sum_e \psi_e'(\|y_e\|^2)\|y_e\|^2.$$

Since $\psi_e'(\|y_e\|^2) > 0$ unless $\|y_e\|^2 \geqslant D_e$ or $\|y_e\|^2 = 0$, this sum cannot be zero unless one of these conditions holds for every edge. □

This proof holds as well if $D_e = \infty$, in which case $L_{\mathcal{F}}^{\Phi}$ vanishes precisely on $H^0(G;\mathcal{F})$. As a result, we have a large class of local nonlinear dynamics on $C^0(G;\mathcal{F})$ that converge to sections of $\mathcal{F}$, given by any set of strictly increasing functions ψ_e. That is, the nonlinear heat equation

$$\dot{x} = -L_{\mathcal{F}}^{\Phi}(x)$$

has trajectories that converge to global sections of $\mathcal{F}$. Further, because $L_{\mathcal{F}}^{\Phi}(x) \in \operatorname{im}\delta^* \perp \ker\delta$ for every x, these trajectories still preserve the projection of the initial condition onto $H^0(G;\mathcal{F})$, and hence converge to the nearest global section to the initial condition.

Returning to the case of finite D_e, there are additional stationary points. In a neighborhood of one of these stationary points, the nonlinear heat equation behaves like a standard heat equation over a different sheaf. For $x \in C^0(G;\mathcal{F})$, define $\mathcal{F}_x$

to be the quotient sheaf of $\mathcal{F}$ with all stalks over edges e with $\|\delta_e x\|^2 \geqslant D_e$ set to zero. If $\|\delta_e x\|^2 \neq D_e$ for all e, there is a neighborhood U of x for which $L^\phi_{\mathcal{F}}(y) = L^\phi_{\mathcal{F}_x}(y)$ for all $y \in U$. We can think of these thresholds as partitioning $C^0(G;\mathcal{F})$ into domains on which consensus dynamics for different sheaves operate. Of course, these domains are not isolated—the dynamics can transport the state from one region to another. However, on a sufficiently small region around a generic equilibrium, the dynamics stays in its initial region. In other words, most equilibria are Lyapunov stable, meaning that trajectories have small perturbations for sufficiently small perturbations of the initial state.

Proposition 7.3.2. *Let $L^\phi_{\mathcal{F}}(x) = 0$, with $\|\delta_e x\|^2 \neq D_e$ for all e. Then x is Lyapunov stable as a fixed point of the threshold heat equation $\dot{x} = -L^\phi_{\mathcal{F}}(x)$. That is, there exist arbitrarily small neighborhoods U, U' of x such that all trajectories beginning in U' stay in U and converge to a section of $\mathcal{F}_x$.*

Proof. Since x is not on the boundary of the region on which the dynamics are given by $L^\phi_{\mathcal{F}_x}$, the center manifold passing through x is precisely the space $H^0(G;\mathcal{F}_x)$ of global sections of $\mathcal{F}_x$. Further, linearizing at x shows that the only non-central manifold is stable. The center manifold theorem [Kel67] then implies that stability of x is equivalent to its stability in the center manifold. But the dynamics are trivial in the center manifold—they are simply constant. Therefore x is Lyapunov stable. □

7.3.2 *Other potential functions*

Another simple choice for ϕ is given by a 1-cochain of $\mathcal{F}$. If $y \in C^1(G;\mathcal{F})$, we can define $\phi_e(x_e) = \|x_e - y_e\|^2$. Then $\nabla\phi(\delta x) = \delta x - y$, which is zero if and only if $\delta x = y$. Further, $L^\phi_{\mathcal{F}}(x) = 0$ if and only if $(\delta x - y) \perp \operatorname{im}\delta$, and if $y \in \operatorname{im}\delta$ this can only happen if $\delta x = y$, so the stationary points of the nonlinear heat equation are precisely the cochains x with $\delta x = y$. If $y \notin \operatorname{im}\delta$, the stationary points are those such that δx is equal to the orthogonal projection of y onto $\operatorname{im}\delta$, i. e.the minimizers of $\|\delta x - y\|^2$. Equilibria of this nonlinear heat equation form an affine subspace of $C^0(G;\mathcal{F})$ corresponding to a shifted copy of $H^0(G;\mathcal{F})$. We can therefore use these Laplacians to solve distributed least squares problems for the sheaf coboundary operator.

Another choice for ϕ is to choose some collection of edges E_N of G, and let $\phi_e = -\|\delta_e x\|^2$ for $e \in E_N$ and $\phi_e = \|\delta_e x\|^2$ for $e \notin E_N$. The corresponding operator $L^\phi_{\mathcal{F}}$ is linear, and is equal to $\delta^* S \delta$, where $S_e = -\operatorname{id}_{\mathcal{F}(e)}$ for $e \in E_N$ and $S_e = \operatorname{id}_{\mathcal{F}e}$ for $e \notin E_N$. Sections of $\mathcal{F}$ are in the kernel of $L^\phi_{\mathcal{F}}$, but this matrix is no longer necessarily positive semidefinite. The corresponding heat equation can be interpreted as a competition between edges in E_N, which seek to increase the

discrepancy of x on their endpoints, and edges not in E_N, which seek to make the values of x at their endpoints more consistent.

Suppose E_N is a cutset of G, dividing G into G_1 and G_2. If the natural map $H^0(G, G_2; \mathcal{F}) \to H^0(G_1; \mathcal{F})$ is not surjective, $L_{\mathcal{F}}^{\phi}$ is indefinite. To see this, take some nontrivial $x \in H^0(G_1; \mathcal{F})$ which is not in the image of this map. Then, extending x by zero to the rest of G, we see that $\delta_e x \neq 0$ for some $e \in E_N$, while $\delta_e x$ vanishes for any $e \notin E_N$. Therefore, $\langle x, L_{\mathcal{F}}^{\phi} x \rangle = \sum_{e \in E_N} \langle \delta_e x, -\delta_e x \rangle$, which is negative because there is at least one nonzero term.

7.4 RANDOM WALKS

The heat equation on $\mathbb{R}^n$ is closely connected with the behavior of random walks, since it defines the evolution of a probability distribution for the location of the random walker. Similarly, the heat equation on a graph describes the evolution of the probability distribution for a continuous-time random walk process on the graph, and the adjacency matrix describes the evolution of a discrete-time random walk. These interpretations do not directly extend to the heat equation on a sheaf.

Discrete-time random walks on $O(n)$-bundles have been considered previously [CZ12; Kem15]. Essentially, a random walk is run on the underlying graph, and an element in the fiber follows this walk via parallel transport. Thus, the state of the random walk at any given point in time is a vector $x[t] = x_v \in \mathcal{F}(v) = \mathbb{R}^n$. At each step, a neighbor u of v is chosen at random, and the next state is then $\mathcal{F}_{u \trianglelefteq e}^{-1} \mathcal{F}_{v \trianglelefteq e} x_v \in \mathcal{F}(u)$. The normalized adjacency matrix $AD^{-1} = I - LD^{-1}$ represents this process, although it no longer gives probability distributions for the random walk. Rather, we have $\mathbb{E}[x[t+1]] = AD^{-1}\mathbb{E}[x[t]]$.

Of course, there is no need to restrict specifically to $O(n)$-bundles in this definition; we can similarly define a random walk for any discrete vector bundle on a graph. The relevant evolution matrix is no longer the normalized adjacency matrix, however, since restriction maps are not unitary. Rather, we have a *walk matrix* $W_{\mathcal{F}}$ with $W_{\mathcal{F}}[u, v] = d_v^{-1} \mathcal{F}_{u \trianglelefteq e}^{-1} \mathcal{F}_{v \trianglelefteq e}$. The degree d_v here is scalar, given by the underlying graph. Again, $\mathbb{E}[x[t+1]] = W_{\mathcal{F}}\mathbb{E}[x[t]]$. If $D^{-1}x$ is a section of $\mathcal{F}$, then x is a stationary distribution for this random walk.

Indeed, this construction extends a bit further. Suppose we have a cosheaf $\hat{\mathcal{F}}$ on a graph, where all extension maps are surjective. We define a similar random walk on $\hat{\mathcal{F}}$, driven by a random walk in the underlying graph. The difference, however, is that parallel transport of a vector over an edge is no longer uniquely defined, which gives an additional entry point for randomness. We need to decide how to pull an element of $\hat{\mathcal{F}}(v)$ back to an element of $\hat{\mathcal{F}}(e)$ for $v \trianglelefteq e$. We do this by specifying a probability distribution $P_{v \trianglelefteq e}$ on $\ker \hat{\mathcal{F}}_{v \trianglelefteq e} \subseteq \hat{\mathcal{F}}(e)$ for every incident pair. Then, we pull $x_v \in \hat{\mathcal{F}}(v)$ back to $\hat{\mathcal{F}}(e)$ by adding an independent random variable $Y_{v \trianglelefteq e}$

distributed as $P_{v \trianglelefteq e}$ to the least-squares solution x_e to the problem $\hat{\mathcal{F}}_{v \trianglelefteq e} x_e = x_v$. We then push down to the other end of the edge via the map $\hat{\mathcal{F}}_{u \trianglelefteq e}$. Thus, the next iteration has $x[t+1] = \hat{\mathcal{F}}_{u \trianglelefteq e}(\hat{\mathcal{F}}^{\dagger}_{v \trianglelefteq e} x_v + Y_{v \trianglelefteq e})$. If we choose these distributions so that $\mathbb{E}[Y_{v \trianglelefteq e}] = 0$, the evolution of the expectation is again described by a walk matrix $W_{\hat{\mathcal{F}}}$ with $W_{\hat{\mathcal{F}}}[u, v] = d_v^{-1} \hat{\mathcal{F}}_{u \trianglelefteq e} \hat{\mathcal{F}}^{\dagger}_{v \trianglelefteq e}$. If $D^{-1}x$ is a harmonic 0-chain of $\hat{\mathcal{F}}$, x is again a stationary distribution under the random walk.

Suppose now that $f : H \to G$ is a graph morphism where all vertices map to vertices and all edges map to edges. Consider the pushforward cosheaf $f_* \overline{\mathbb{R}}$, where $\overline{\mathbb{R}}$ is the constant cosheaf on H. We can therefore interpret the elements of $C_0(G; f_* \overline{\mathbb{R}})$ as describing functions on the vertices of H, and in particular, there is some subset corresponding to probability distributions on the vertices of H. What happens if we run a random walk on $f_* \overline{\mathbb{R}}$ begun at $f_* \mathbb{1}_v$, for some $v \in H$? At the first step the state is $f_* \mathbb{1}_v \in f_* \overline{\mathbb{R}}(f(v))$. We will choose uniformly at random a neighbor of $f(v)$, and then pull back to $f_* \overline{\mathbb{R}}(e)$, where e is an edge in G incident to $f(v)$. This space has a natural basis that includes representatives for a number of edges incident to v in H, and the preimage of $f_* \mathbb{1}_v$ lies within the subspace spanned by these representatives. The least-squares preimage distributes the weight of $f_* \mathbb{1}_v$ equally across these representatives. Thus, if there are d_{ve} edges e_i in H incident to v lying over $e \in G$, $f_* \mathbb{1}_v$ pulls back to $d_{ve}^{-1} \sum f_* \mathbb{1}_{e_i} \in f_* \overline{\mathbb{R}}(e)$. If we let $Y_{f(v) \trianglelefteq e} = 0$ be deterministic, we can interpret the pulled-back cochain as a probability distribution on the fiber of e. The map $f_* \overline{\mathbb{R}}_{u \trianglelefteq e}$ preserves probability distributions, so we can interpret the next step as not just a value in $f_* \overline{\mathbb{R}}(u)$, but as a probability distribution on $f^{-1}(u)$. This interpretation is preserved over subsequent steps, so the random walk on $f_* \overline{\mathbb{R}}$ determines a random walk on H, driven by a random walk on G. In general this is not the same as the uniform random walk on H, since in essence an edge of G is first chosen and then an edge in H lying over that edge is chosen, which means that neighbors are not uniformly selected in H.

In general, sheaf Laplacians will not have much to say about these sorts of holonomic random walks, because their evolution is governed by the walk matrix $W_{\hat{\mathcal{F}}}$. Only in the case of $O(n)$-bundles are these equivalent, and then the walk matrix is equal to $I - LD^{-1}$.

Part II

APPLICATIONS

8

DISTRIBUTED COORDINATION

8.1 DISTRIBUTED CONSENSUS

The problem of getting a network of agents to agree on some data or policy without centralized coordination has been a subject of intense interest for the past several decades. This *consensus problem* can be addressed with methods from spectral graph theory and control theory. A core tool is *distributed averaging*: agents take a series of steps averaging their local values with those of their neighbors. Given edge weights A_{ij}, the agents update their state via

$$x_i[t+1] = \sum_{i \sim j} A_{ij} x_j[t].$$

The properties of adjacency matrices of graphs are obviously important for understanding these distributed averaging algorithms [OM04; Boy+05; OT11]. Complications can be larded atop this model, including time-varying networks, time delays, and noisy communication, and this general framework finds applications in many fields [JLM03; GS10; Bul19].

Graph Laplacians are also a useful tool for the analysis and construction of such algorithms. As was noted in Chapter 7, trajectories of the heat equation $\dot{x} = -Lx$ corresponding to a graph Laplacian converge to the projection of the initial condition onto the kernel of L. This kernel is spanned by the constant vector $\mathbb{1}$, so in fact the trajectories converge to a global consensus on the average of the initial values. The Laplacian flow is therefore a continuous-time distributed averaging process.

The sheaf Laplacian flow suggests an immediate generalization to a setting where the desired distributed state is not a constant value, but rather a section of a sheaf. In the simplest case, this might just be a particular set of ratios between values at different vertices, realized by the space of sections of an appropriate line bundle. Here the kernel of $L_{\mathcal{F}}$ is spanned by a single non-constant vector. However, the sheaf formulation extends to significantly more complex subspaces, like the space of trajectories of a distributed linear system as in Chapter 10.2.1. If a vector space can be realized as the space of sections of a sheaf on an appropriate graph, the sheaf Laplacian flow gives an algorithm for distributed consensus on this vector space.

The problem of distributed algorithms for a graph G that converge to projections onto a given subspace of $\mathbb{R}^{V(G)}$ has previously been approached from a direct discrete-time perspective [BSB09; LBS19; NVS19]. Typically the subspace considered is a space spanned by eigenvectors of the graph Laplacian with small eigenvalues—in graph signal processing terms, a subspace of low-frequency signals. The sorts of problems addressed by these methods are mostly orthogonal to those which can be addressed with sheaf Laplacians. The space of global sections of a sheaf on a graph is typically uninteresting unless vertex stalks have dimension greater than 1. For a sheaf with all vertex stalks 1-dimensional, H^0 can only have dimension greater than 1 if there are localized sections; one cannot construct a sheaf on G whose sections are the space of low-frequency vertex signals on G.

To produce an implementable algorithm from the sheaf heat equation, we need to have a discrete-time version of the dynamics. The easiest way to do this is to use an Euler-type discretization:

$$x[t+1] = x[t] - \alpha L_{\mathcal{F}} x[t].$$

In essence, we use the matrix $T_{\mathcal{F}} = I - \alpha L_{\mathcal{F}}$ as a linear update rule for the state. We need to choose α properly to ensure convergence. Every eigenvalue of $T_{\mathcal{F}}$ is of the form $1 - \alpha\lambda$ for some eigenvalue λ of $L_{\mathcal{F}}$. For the discrete-time iteration to be stable, we need to ensure that these eigenvalues all lie in $(-1, 1]$. If λ_2 is the smallest nonzero eigenvalue of $L_{\mathcal{F}}$ and λ_{max} is the largest eigenvalue of $L_{\mathcal{F}}$, we therefore need $\alpha < 2/\lambda_{max}$. The convergence rate is determined by the nontrivial eigenvalue farthest from 0. Depending on α, this may be either $1 - \alpha\lambda_{max}$ or $1 - \alpha\lambda_2$. The choice of α optimizing the convergence rate is $\alpha = \frac{2}{\lambda_2 + \lambda_{max}}$, giving a maximal eigenvalue of $\frac{\lambda_{max} - \lambda_2}{\lambda_{max} + \lambda_2}$.

8.1.1 *Optimizing the convergence rate*

The inner products on edge stalks of $\mathcal{F}$ are required to construct the Laplacian, but they do not affect the long-term limiting behavior of the heat equation. We are therefore free to modify them in order to influence shorter-term behavior like the convergence rate. For graph Laplacian consensus, doing this amounts to choosing edge weights to optimize convergence speed, which has been studied by Xiao and Boyd [XB04] as a semidefinite program. Other formulations apply to natively discrete-time consensus algorithms [OT11].

A slight generalization of Xiao and Boyd's approach lets us optimize the convergence rate for sheaf consensus. Given a matrix representation B of the coboundary operator δ, we can change the effective inner product on C^1 by choosing a block diagonal positive semidefinite matrix W and using the Laplacian $L_{\mathcal{F}}^W = B^T W B$.

In essence, we replace the standard inner product on $C^1(G;\mathcal{F})$ with the inner product $x^T W x$. In order to optimize the convergence rate, we want to keep the smallest and largest nontrivial eigenvalues of $B^T W B$ as close to 1 as possible. The largest eigenvalue can be controlled by a semidefinite bound $B^T W B \preceq \lambda I$, but λ_2 is a bit trickier to control. Let Π_{H^0} be the orthogonal projection onto $H^0(G;\mathcal{F})$. If $\lambda(I - \Pi_{H^0}) \preceq B^T W B$, then $\lambda_2 \geqslant \lambda$, since this controls the spectrum of $B^T W B$ on the orthogonal complement of $H^0(G;\mathcal{F})$. A bit of reformulation gives us the following semidefinite program:

$$\begin{aligned} \min_{t, W_e} \ & t \\ \text{s.t. } & W_e \succeq 0 \quad \forall e \\ & (1+t)I - B^T \operatorname{diag}(W_e) B \succeq 0 \\ & B^T \operatorname{diag}(W_e) B + \Pi_{H^0} - (1-t)I \succeq 0 \end{aligned}$$

8.2 FLOCKING

Here is a simple example of a nontrivial sheaf useful in a plausible coordination problem. Consider a collection of autonomous agents in $\mathbb{R}^3$, each of which has its own coordinate system with respect to which it measures the world. Agents can calculate the bearings to other nearby agents in the form of unit vectors represented in their own reference frames, but know nothing about their neighbors' reference frames. The problem is to have the agents agree on a consensus direction in the global frame, represented in their own local reference frames. Such a consensus would be useful perhaps in order for the agents to agree on a common direction of travel.

The data of this problem forms a sheaf $\mathcal{F}$ on the neighborhood graph G of the agents. The vertex stalks are $\mathcal{F}(v) = \mathbb{R}^3$, containing a vector in the local frame, while the edge stalks are $\mathcal{F}(e) = \mathbb{R}^1$, representing the direction along the bearing between the agents at either end of the edge. Let $b_{uv} \in \mathbb{R}^3$ be the unit vector pointing from u to v, represented in u's reference frame. For an oriented edge $e = u \sim v$, the restriction map $\mathcal{F}_{u \trianglelefteq e}$ is given by the matrix b_{uv}^T, while the restriction map on the other end side is given by $-b_{vu}^T$. (Note that this requires the agents to agree on an orientation. One way to ensure this is if every agent has a unique serial number, in which case the ordering of numbers can induce an orientation.)

What are the sections of $\mathcal{F}$? Note that by passing to a unitarily isomorphic sheaf, we may assume that all the local reference frames are in fact equal. In this situation we hope that only the constant cochains of $\mathcal{F}$ would be sections, since these would

correspond exactly to globally consistent directions. However, this is not quite true: there are a few extra sections to reckon with.

It is perhaps easier to understand $H^0(G;\mathcal{F})$ by considering the dual cosheaf $\hat{\mathcal{F}}$. We can interpret this cosheaf as a subcosheaf of the constant cosheaf $\overline{\mathbb{R}^3}$ on G. Elements of vertex stalks of the cosheaf correspond to *forces* exerted on vertices of G, while elements of the edge stalks correspond to *stresses* experienced along the lines connecting the vertices. The boundary map $\partial : C_1(G;\hat{\mathcal{F}}) \to C_0(G;\hat{\mathcal{F}})$ takes an assignment of stresses to each edge of G and gives the net force it effects on each vertex of G. The cosheaf homology $H_0(G;\hat{\mathcal{F}}) = C_0(G;\hat{\mathcal{F}})/\operatorname{im}\partial$ is isomorphic to $H^0(G;\mathcal{F})$, and represents systems of net forces on the vertices of G that cannot be imposed by a system of stresses acting along the edges of G. There is always at least a six-dimensional space of such forces: those that generate uniform translations of the vertices and those that generate rotations of G. Thus $H^0(G;\mathcal{F})$ contains the constant cochains, but also a three-dimensional space of spurious non-constant sections. In other words, $\mathcal{F}$ is not quite an approximation to the constant sheaf $\underline{\mathbb{R}^3}$.

One way to overcome this difficulty is to note that the space of nonconstant sections of $\mathcal{F}$ depends on the configuration of the agents. If the agents change their relative positions in a sufficiently generic way, giving a new sheaf $\mathcal{F}^1$, any nonconstant section of $\mathcal{F}$ will not be a section of $\mathcal{F}^1$. Indeed, given a section x of $\mathcal{F}$, all that is necessary is for each agent v to travel a small distance in the direction x_v. This changes the bearing vector b_{uv} by a vector proportional to $x_v - x_u$, and hence changes the restriction maps by the same amount. Thus x is also a section of $\mathcal{F}^1$ if and only if $(x_v - x_u)^T x_v - (x_v - x_u)^T x_u = 0$, which happens only if $x_v = x_u$, i. e.x is a constant section.

8.3 EFFICIENT COMMUNICATION

A network running a consensus algorithm of the sort we have been discussing requires a significant amount of communication between nodes. At each time step, a pair of nodes u, v connected by an edge e must exchange messages containing $\mathcal{F}_{u \trianglelefteq e} x_u$ and $\mathcal{F}_{v \trianglelefteq e} x_v$, i.e two vectors in $\mathcal{F}(e)$. The amount of communication required at each step is thus proportional to $\dim C^1(G;\mathcal{F})$. In the case of consensus with respect to the constant sheaf on a graph, this communication cost is proportional to the number of edges in the graph. This has made graph sparsification an appealing tool for improving the performance of consensus algorithms. The total time taken to converge to within ϵ of a section depends on the spectrum of L as well as the amount of communication required for each step. If μ is the spectral radius of the discrete-time consensus operator and D is the amount of communication required at each step, the convergence rate is $O(\mu^{n/D})$, with n the

number of communication steps. A sparsified graph might reduce D substantially while increasing μ only slightly, leading to an improved overall convergence rate.

For an example, consider consensus on a complete graph, with $O(n_v^2)$ edges and dominant eigenvalue μ. We can sparsify this to a graph with $O(n_v/\epsilon^2)$ edges with dominant eigenvalue $(1+\epsilon)\mu$. We wish to compare $\mu^{1/O(n_v^2)}$ with $((1+\epsilon)\mu)^{1/O(n_v/\epsilon^2)}$, or equivalently, two quantities that grow as $\frac{\log \mu}{n_v^2}$ and $\frac{\epsilon^2}{n_v}(\log \mu + \log(1+\epsilon))$. It is easy to see that the second quantity grows more slowly, so for sufficiently large graphs, it is preferable to use a sparsified graph rather than the all-to-all network.

The same analysis holds for using a spectrally good approximation to $\mathcal{F}$ to run a generalized consensus algorithm using the sheaf Laplacian. Reducing the dimension of $C^1(G;\mathcal{F})$ while keeping the spectrum of $L_{\mathcal{F}}$ tightly concentrated can improve the overall convergence rate. Applying this to the constant sheaf $\underline{\mathbb{R}^n}$, we are brought to the intriguing idea that an approximation to the constant sheaf might improve communication efficiency even more than a sparsified graph.

Consider, for instance, an expander sheaf of algebraic degree k on a graph of degree ℓ, and compare this with a constant sheaf on a graph of degree k. The expander sheaf has adjacency spectral radius at least $2\frac{k}{\ell}\sqrt{\ell-1}$ by the discussion in Chapter 6.2.1, while the Alon-Boppana bound on the spectral radius of the graph is $2\sqrt{k-1}$. For $k > 2$ the expander sheaf bound is tighter. An equivalent inequality is

$$\frac{\sqrt{\ell-1}}{\ell} \leqslant \frac{\sqrt{k-1}}{k}$$

for $2 < k \leqslant \ell$; equivalently, the function $x \mapsto \frac{\sqrt{x-1}}{x}$ is decreasing for $x \geqslant 3$. The derivative of this function is $\frac{1-\frac{x}{2}}{x^2\sqrt{x-1}}$, which is negative for $x > 2$. The interpretation of this fact in terms of communication efficiency for consensus networks is that for a given communication budget per step, an expander sheaf has the possibility of giving faster consensus than the constant sheaf on an expander graph.

Thus far we have glossed over precisely what the cost of communication is for a step of the consensus algorithm, since this depends on the precise implementation of the communication structure for the network. Counting the communication cost proportional to $C^1(G;\mathcal{F})$ implicitly assumes that only one communication link in the graph can be active at a time, which is not typically a very realistic model. In most cases it is more reasonable to assume that either all links may be active at once or that each node can only communicate over one link at a time. In the first case, the communication cost is then proportional to $\max_e \dim \mathcal{F}(e)$, and in the second case, it is bounded below by $\max_v \sum_{v \trianglelefteq e} \dim \mathcal{F}(e)$. This suggests that merely controlling the dimension of $C^1(G;\mathcal{F})$ in an approximation is not sufficient. Rather, we wish to control dimensions of edge stalks locally. As noted in Chapter 5.3, this is a

significantly more difficult problem. A general construction of expander sheaves would be a partial resolution of this problem for approximations to the constant sheaf, since the regularity requirement bounds the local dimension of edge stalks.

The problem of minimizing communication for sheaf Laplacian-based consensus can be seen as a nonconvex generalization of the optimization problem of Section 8.1.1. Rather than simply minimizing the nontrivial spectral radius of $I - L_{\mathcal{F}}$, we minimize the effective convergence rate $\mu^{1/D}$, where D represents the amount of communication cost per step. This amounts to an optimization problem

$$\begin{aligned}
&\min\ t^{1/D} \\
&\text{s.t. } W_e \succeq 0 \quad \forall e \\
&\qquad (1+t)I - B^T \operatorname{diag}(W_e) B \succeq 0 \\
&\qquad B^T \operatorname{diag}(W_e) B + \Pi_{H^0} - (1-t)I \succeq 0 \\
&\qquad D = \sum_e \operatorname{rank} W_e
\end{aligned}$$

which depends on the ranks of the edge weight matrices.

9

NETWORK SCIENCE AND SHEAVES

Network science typically focuses on graphs as models for networks. This allows us to model the patterns of connection between nodes in a network, possibly (in the case of weighted graphs) with some measure of intensity of connection. Sheaves offer the potential for much richer descriptions of network structure. They can encode the nature of relationships between nodes, not just their presence or absence.

9.1 OPINION DYNAMICS

Social networks are one of the canonical examples of networks that arise in the real world. For decades, social scientists have developed models for the way that opinions change as individuals communicate in a social network. Many early investigations were based on the simple idea that individuals linked in a social network tend to adjust their opinions toward consensus. This led to local averaging dynamics based on the Laplacian or adjacency matrices [Tay68; DeG74].

More sophisticated nonlinear models include the popular *bounded confidence* dynamics. This model is similar to the local averaging model, but individuals have a threshold of disagreement above which they do not take their neighbors' opinions into account. One famous implementation of this idea was popularized by Dittmer and Hegselmann-Krause [Dit01; HK02]. Given a weighted adjacency matrix A, the discrete-time evolution of an opinion distribution x proceeds as

$$x_i[t+1] = \sum_j A_{ij} \mathbb{1}_{|x_i[t]-x_j[t]| \leqslant \delta_{ij}} x_j[t].$$

This model has, at some difficulty, been extended to continuous time while maintaining the discontinuous confidence thresholds [CF11].

Models of opinion dynamics have usually focused on single-dimensional opinions, those that can be represented by a single parameter per individual in the network. When higher-dimensional opinions have been considered, these have typically been coupled purely independently.

Cellular sheaves offer a new way to construct principled models of opinion networks. The communication structure within a social network can be modeled

with a *discourse sheaf* on a graph, where vertex stalks represent spaces of private opinions and edge stalks represent spaces in which discussion happens; restriction maps then translate between privately-held opinions and public discussion. The discourse sheaf specifies when individuals consider themselves to be in agreement: sections of the sheaf are precisely those distributions of private opinions that agree when mapped to the pairwise discourse spaces.

A discourse sheaf can model various new phenomena in social networks. These include

- Discussing only a subset of one's opinions with a given neighbor. This is implemented by setting the relevant restriction map to projection onto a subspace spanned by a set of discussed topics.
- Generating opinions on specific policies from privately-held general principles. Here the discourse space over an edge is spanned by the policies under consideration, while the vertex stalk is spanned by the space of principles. The restriction maps of the sheaf implement the translation from principles to policies.
- Modulating the intensity or even the direction of one's opinions. Restriction maps can scale or change the sign of opinions as they are translated to the discourse space. This can allow individuals with opinions far from the mainstream to appear to be in agreement with the rest of the network.

Once we have specified the discourse sheaf, we can begin to study the dynamics of opinions living in this sheaf. The dynamical systems studied in Chapter 7 come in handy here. We can of course begin with the sheaf diffusion dynamics $\dot{x} = -L_{\mathcal{F}}x$. This will converge to an opinion distribution with agreement in every pairwise discourse space. Diffusion dynamics are only the beginning. Suppose there is a collection of *stubborn* individuals who refuse to change their opinions, while the rest of the population proceeds according to the diffusion dynamics. This amounts to the heat equation with boundary conditions discussed in Chapter 7.1.1. In particular, the global opinion distribution converges to a harmonic extension of the opinions of the stubborn individuals.

Most of the other dynamical systems discussed in Chapter 7 have a natural interpretation in terms of discourse sheaves. The control system dynamics (7.1) describe the manipulation of opinion distributions by feeding information to a subset of individuals. Proposition 7.1.2 gives conditions under which it is possible to steer the opinion distribution to any section of the discourse sheaf.

Dynamics on the sheaf structure can be interpreted as agents adapting their strategies for communicating with their neighbors. Given an initial opinion distribution, the restriction map dynamics (7.3) converge to a communication strategy

that eliminates disagreement between neighbors. One might say that the agents "learn to lie" about their true opinions in order to keep the peace. The combined heat equation and restriction map dynamics (7.5) involve individuals adapting both their opinions and the way they express them to their neighbors, leading ultimately to a stable discourse sheaf/opinion distribution pair.

The threshold dynamics described in Section 7.3.1 are a continuous-time smoothed version of bounded-confidence opinion dynamics. Rather than a sharp threshold of disagreement over which neighbors ignore each other, opinion influence is attenuated as disagreement increases, diminishing to zero above the threshold. Proposition 7.3.1 implies that we can think of this as dynamics with a sheaf structure that varies depending on the opinion distribution, and all equilibria are consistent over edges whose disagreement is below the threshold. These equilibria are Lyapunov stable by Proposition 7.3.2.

There are two ways to encode signed social relationships. One is the $O(1)$-bundle approach discussed in Chapter 2.1. The edge between two friends has both restriction maps the same sign, while the edge between two enemies has their signs opposite.[1] When the diffusion dynamics are applied to cochains on such a sheaf, we see that enemies attempt to adjust their opinions until they have opposite signs and the same magnitude. Alternately, we may use the signed Laplacian from Chapter 7.3.2 given by a negated different edge potential function over edges between enemies. Under these dynamics, enemies push their opinions to be as far from agreement as possible. This means that for certain patterns of friendship and foeship, the opinion dynamics are unstable, leading to unbounded opinion distributions.

9.2 LEARNING SHEAF LAPLACIANS

To do network science, it is first necessary to have a network. In many situations, the edges of a network are given explicitly: many physical networks, social networks, and hypertext networks, for instance. However, in other settings, edges must be reconstructed from lower-order data. This is particularly common in network approaches to physical and natural systems, where, for example, it is difficult to observe the entire collection of relationships between a set of genes or neurons or species. In neuroscience, this sort of analysis is known as extracting the *functional connectivity* from data associated to nodes.

Approaches to constructing functional networks from data are often quite simple. For instance, a very common method is to compute a correlation matrix between different data sources and impute connections corresponding to entries whose absolute value exceed some threshold. These correlation networks can be

1 Under the semantics we have been using, this would mean that one enemy lies about their opinion.

made more sophisticated, with weights and factorization analyses, but they still typically amount to extracting only connections without information about relationships [Hor11]. We want to do better, and learn a sheaf structure, not just a network structure, from data.

To do this, we will adapt a framework from the field of graph signal processing [Ort+18]. The central concept of this field is that we can treat data associated with graphs in ways similar to those used for traditional signal processing. This involves the construction of Fourier-like transforms using Laplacian and adjacency eigenvectors, filtering and sampling theory, and learning from graph signals. In particular, the interpretation of the graph Laplacian as giving a measure of the *smoothness* of a signal is central. The Laplacian quadratic form $E(x) = \langle x, Lx \rangle$ is a measure of the ℓ^2 variation of a signal over the edges of a graph. A smooth signal has low variation; the smoothest signals are the constant functions.

In many situations, we expect data that comes from a process on a graph to be smooth with respect to that graph's Laplacian. This fact has been leveraged to build frameworks for learning a graph from data [LT10; Don+15; Kal16; EPO17]. The idea is this: construct an optimization problem over an appropriate space of Laplacian (or Laplacian-like) matrices, where the cost function contains a term measuring the smoothness of the observed data with respect to the proposed Laplacian. That is, if we have a set of functions x^k on the vertices of our graph, the cost function is

$$C(L) = \sum_i \langle x^k, Lx^k \rangle + \text{regularization terms.}$$

The regularization terms typically include a term to control sparsity of L, as well as a term that breaks the homogeneity of the cost function, so that L does not go to zero and represents a connected graph. For instance, Kalofolias [Kal16] proposed the cost function

$$C(L) = \sum_k \langle x^k, Lx^k \rangle - \alpha \sum_i \log(L_{ii}) + \beta \sum_{i<j} \left|L_{ij}\right|^2, \tag{9.1}$$

while Egilmez et al. [EPO17] proposed functions of the form

$$C(L) = \sum_k \langle x^k, Lx^k \rangle - \log\det L + \alpha \sum_{i<j} \left|L_{ij}\right|. \tag{9.2}$$

These cost functions are minimized over the cone of graph Laplacians, or perhaps over some larger cone, such as the set of positive definite graph Laplacians with some signs changed, or the cone of diagonally dominant matrices. Since these are convex sets with simple descriptions, it is typically possible to formulate these problems in a way that allows efficient computation. For instance, the cost function (9.1)

can be equivalently written in terms of the adjacency matrix with nonnegative constraints and solved via a simple primal-dual algorithm. The cost function (9.2) does not admit such a reformulation, but remains a convex optimization problem that can be expressed as a semidefinite program.

This smoothness-based approached to learning a graph Laplacian generalizes readily to learning a sheaf Laplacian. We need to construct appropriate cost functions for sheaf Laplacians and optimize them over the relevant spaces of sheaf Laplacians. For tractability and optimality guarantees, we focus on convex objective functions. To begin, we assume that the dimension of stalks over vertices of the graph are known, while the dimension of edge stalks is allowed to be arbitrary. That is, we know how to partition our input signals x^i into pieces that come from each node. A generalization of (9.1) to sheaf Laplacians is

$$C(L) = \sum_k \langle x^k, Lx^k \rangle - \alpha \sum_i \log \operatorname{tr} L_{ii} + \beta \sum_{i<j} \|L_{ij}\|_F^2. \tag{9.3}$$

This is a sum of convex functions, and hence convex. The second term ensures that the diagonal blocks of L do not go to zero, while the third term exerts some control over the sparsity of the off-diagonal blocks of L—as β increases it encourages weight to be distributed more evenly across these blocks.

Similarly, a generalization of (9.2) to sheaf Laplacians is

$$C(L) = \sum_k \langle x^k, Lx^k \rangle - \log\det L + \beta \sum_{i<j} \|L_{ij}\|_1, \tag{9.4}$$

where $\|L_{ij}\|_1$ is the entrywise ℓ^1 norm of the matrix L_{ij}, not the operator ℓ^1 norm.

The structure of various cones of sheaf Laplacians is described in Chapter 3.3.2. This representation allows us to express our optimization problem in a format suitable for implementation with standard convex solvers. For instance, if we wish to optimize $C(L)$ over the cone of all sheaf Laplacians with given vertex stalk dimensions, we solve the problem

$$\begin{aligned} \min_L \ & C(L) \\ \text{s.t. } & L = \sum_e L_e \\ & L_e \succeq 0 \\ & L_e \text{ has the correct sparsity pattern} \end{aligned} \tag{9.5}$$

To optimize over the cone of matrix-weighted graphs, we solve

$$\begin{aligned} \min_{L} \; & C(L) \\ \text{s.t. } & L_{ij} = -W_{ij}, \quad i \neq j \\ & W_{ij} \succeq 0 \\ & \sum_j L_{ij} = 0 \quad \forall i \end{aligned} \tag{9.6}$$

Finally, to optimize over the convex hull of the space of connection Laplacians, we solve

$$\begin{aligned} \min_{L} \; & C(L) \\ \text{s.t. } & L_{ii} = d_i I \quad \forall i \\ & d_i \geqslant \sum_{i \neq j} \|L_{ij}\| \quad \forall i \\ & d = \sum_{ij} w_{ij} \mathbb{1}_{\{i,j\}} \end{aligned} \tag{9.7}$$

The cost function (9.3) can be written in terms of the edge contribution matrices L_e. First concatentate the vertex signals x^k into a matrix X, and note that the smoothness term is equal to $\operatorname{tr}(X^T L X) = \operatorname{tr}(L X X^T)$. Let $M = XX^T$ and construct the matrices M_e with the corresponding sparsity patterns to the matrices L_e. Then $\operatorname{tr}(LM) = \sum_e \operatorname{tr}(L_e M_e)$. Similarly, the off-diagonal regularization term is simply $\sum_e \|\operatorname{offdiag}(L_e)\|_F^2$. Only the term $\sum_i \log \operatorname{tr} L_{ii}$ requires us to combine the individual matrices L_e, becoming $\sum_i \log\left(\sum_{i \trianglelefteq e} \operatorname{tr} L_e[i,i]\right)$. Thus, we have

$$C_{\text{sheaf}}(L) = \sum_e \operatorname{tr}(L_e M_e) - \alpha \sum_i \log\left(\sum_{i \trianglelefteq e} \operatorname{tr} L_e[i,i]\right) + \beta \sum_e \|\operatorname{offdiag}(L_e)\|_F^2.$$

The cost function decomposes even more nicely for the matrix-weighted and connection Laplacian cases:

$$C_{\text{MW}}(L) = \sum_{i \neq j} \operatorname{tr}(W_{ij}(M_{ii} + M_{jj} - 2M_{ij})) - \alpha \sum_i \log\left(\sum_j \operatorname{tr}(W_{ij})\right) + \beta \sum_{i>j} \|W_{ij}\|_F^2$$

$$C_{\text{CL}}(L) = \sum_{i \neq j} \operatorname{tr}(w_{ij}(M_{ii} + M_{jj}) - 2L_{ij}M_{ij}) - \alpha \sum_i \log d_i + \beta \sum_{i>j} \|L_{ij}\|_F^2.$$

These simplifications are more difficult to apply to the cost function (9.4), due to the indecomposability of the term $\log\det L$.

9.2.1 *Statistical interpretations*

The cost function (9.2) has an alternate interpretation as a regularized version of a likelihood function for the inverse covariance matrix of a multivariate Gaussian. To see this, suppose that the signals x^k are distributed with distribution function

$$p(x|L) = \sqrt{\frac{\det L}{(2\pi)^d}} e^{-\frac{1}{2}\langle x, Lx\rangle}.$$

This implies that L is the inverse of the covariance matrix of the random variable x. The log-likelihood is

$$\log p(x|L) = \frac{1}{2}\log\det L - d\log 2\pi - \frac{1}{2}\langle x, Lx\rangle,$$

so the log-likelihood of a series of independent observations x^k is

$$\log p(x^1,\dots,x^n|L) = \sum_i \log p(x^k|L) = -\frac{1}{2}\left[\sum_k \langle x^k, Lx^k\rangle + n\log\det L + K\right].$$

We normalize the log-likelihood, drop constant terms, formulate in terms of the sample covariance matrix $M = \frac{1}{n}XX^T$, and add a sparsity-enforcing ℓ^1 norm term to get the convex functional

$$C(M|L) = \operatorname{tr}(LM) - \log\det L + \alpha\sum_{i>j}\left|L_{ij}\right|.$$

Minimizing this objective subject to the constraint that L be positive semidefinite is a problem known as the *graphical lasso* [BGd08; FHT08].

The sparsity structure of the inverse covariance matrix of a multivariate normal random variable implies certain conditional independence properties of the entries of the random vector. The sparsity pattern determines a graph G with vertices corresponding to entries of x. If a set K of nodes forms a separating set—its removal separates the remaining nodes of G into two components G_1 and G_2, we can conclude that $x|_{G_1}$ and $x|_{G_2}$ are independent given $x|_K$. This property makes the collection of random variables $X|_i$ into a *Gaussian Markov random field*, or a *Gaussian graphical model* [RH05]. This conditional independence structure is implied by the factorization of the probability distribution function of x into a product of

terms, each of which depends only on the two vertices incident to a given edge in G. That is,

$$p(x) = \sqrt{\frac{\det L}{(2\pi)^d}} \prod_e e^{-\frac{1}{2}\langle x, L_e x\rangle}.$$

The approach to graph learning proposed by Egilmez et al. adds additional constraints to the graphical lasso, requiring that the inverse covariance matrix have the form of a graph Laplacian. One motivation for this requirement is that it ensures that the matrices L_e in the factorization are positive semidefinite and hence themselves define Gaussian distributions. However, the edge contribution matrices L_e have this form precisely when L is a sheaf Laplacian. In the language of graphical models, such a decomposition is said to define a *pairwise normalizable* Gaussian Markov random field [MJW06].

10

LINEAR SYSTEMS

10.1 LINEAR TIME-INVARIANT SYSTEMS

Traditionally, a discrete-time linear time-invariant control system is specified as a set of linear equations specifying the time evolution of the state and outputs of the system. That is, a linear time-invariant system is given in the form

$$\begin{aligned} x[t+1] &= Ax[t] + Bu[t] \\ y[t] &= Cx[t] + Du[t], \end{aligned} \tag{10.1}$$

where $x[t]$ is the state at time t, $u[t]$ is the input at time t, and $y[t]$ is the output at time t. Admissible trajectories of the system are solutions of this set of equations, and properties of the system are defined in terms of the linear algebraic properties of the equations [Son98].

Cellular sheaves represent structured systems of linear equations, and we can translate the recurrence relation (10.1) into a sheaf over an infinite graph, demonstrated in Figure 10.1. There are three vertices for each time step: one for the input, one for the output, and one to hold the current state and input of the system. Sections of the sheaf correspond to solutions of (10.1), and hence to admissible trajectories of the dynamical system. One can compute the output behavior of the system for a given sequence of inputs by extending a partial section on the input vertices to a complete section; the structure of the restriction maps of the sheaf guarantees that such an extension exists. Conversely, one may try to find controls that produce a given output trajectory by solving the same problem, but such an extension may not exist—its existence depends on the *controllability* of the underlying system. We call this sheaf the *evolution sheaf* of the control system, and denote it $\mathcal{E}$, with its underlying graph E.

At every central vertex, the evolution sheaf has the same algebraic structure, due to the time-invariance of the dynamical system. This allows us to express the data of the dynamical system in a much more compact sheaf form. This is the *template sheaf* of the dynamical system, illustrated in Figure 10.2. We denote it $\mathcal{T}$,

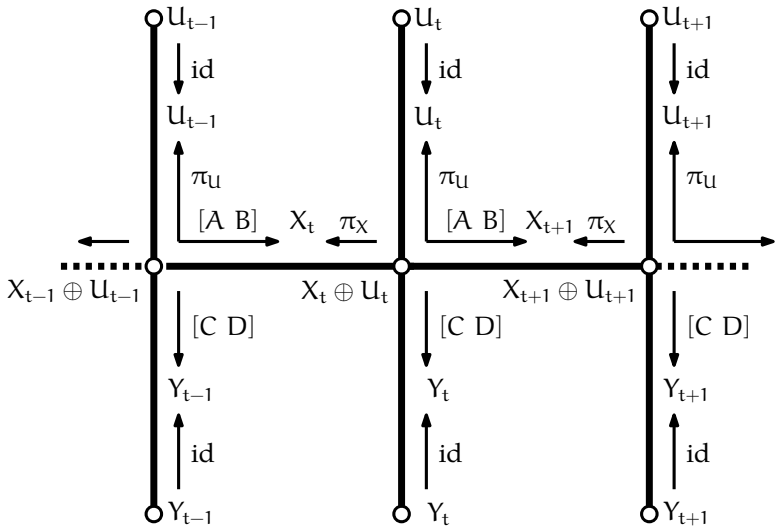


Figure 10.1: The evolution sheaf of a discrete time LTI system.

and its underlying graph T.[1] Note that there is a cell complex morphism $p : E \to T$, sending all central vertices to the vertex of the loop, all vertices labeled U_t to the vertex labeled U, and all vertices labeled Y_t to the vertex labeled Y. The edges between central vertices are mapped to the loop, with the orientation such that incrementing t corresponds to a clockwise path along the loop. We then note that the evolution sheaf is the pullback of the template sheaf, i. e., $\mathcal{E} = p^*\mathcal{T}$.

Since sections of $\mathcal{T}$ pull back to sections of $\mathcal{E}$ that are constant on fibers of p, sections of $\mathcal{T}$ correspond to equilibria of the system. These are not only fixed points in the zero-input regime, but include equilibria of the system with constant nonzero inputs as well.

A reasonable question at this point is when a sheaf over T describes a dynamical system, that is, when does it have the structure of a template sheaf? To recover maps A, B, C, D yielding the standard state-space description of the system, the maps labeled by π_X and π_U must be the structure maps of a product decomposition of the stalk over the central vertex. This, combined with the requirement that the restriction maps from the stalks at the input and output terminals be isomorphisms, ensures that the template sheaf lifts to the evolution sheaf of an LTI system. Given the direct sum decomposition of the stalk at the center, we identify the subspaces X and U, which then determines the four maps A, B, C, and D.

1 This graph is not technically a regular cell complex, since there is a self-loop. We can ameliorate this difficulty by adding an extra vertex, or simply resign ourselves to working slightly outside the formal regularity of the category of regular cell complexes as mentioned in passing in Chapter 1.1.

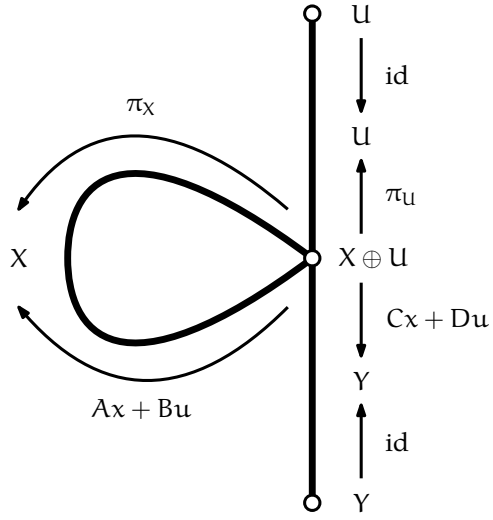


Figure 10.2: The template sheaf of a discrete time LTI system.

What happens if we do not require such a decomposition? We can still view sections of $p^*\mathcal{T}$ as admissible trajectories of an open dynamical system, but this system may not behave in the ways we are accustomed to. For instance, if π_X is not surjective, the system will not be deterministic: multiple trajectories may follow from a given initial state. If the kernels of π_X and π_U do not span their domain, it will not always be possible to change the control input independently of the state, or causality may be violated. In short, sheaves over T represent generalized open discrete-time dynamical systems.

10.2 BEHAVIORAL CONTROL THEORY

These sorts of generalized systems are described in the language of *behavioral control theory*, introduced and championed by Jan Willems [Wil86; Wil07]. Rather than instantiating a control system as a set of differential or difference equations, behavioral control theory conceptualizes it as a relation between possible input and output signals. Indeed, in this formulation, inputs and outputs are formally interchangeable: to be more general, systems simply impose constraints on the values at their boundaries. This is a representation nearer to our understanding of natural laws. After all, the equations of physics do not have inputs or outputs.

To be precise, a discrete-time behavioral control system is a pair $(W, \mathcal{B})$, where W is a state space and $\mathcal{B}$, the *behavior*, is a subset of $W^{\mathbb{Z}}$ consisting of admissible trajectories of the system. Such a system is *linear* if W is a vector space and $\mathcal{B}$ is a subspace of $W^{\mathbb{Z}}$. It is *time invariant* if $T_k\mathcal{B} \subseteq B$ for all k, where $T_k : W^{\mathbb{Z}} \to W^{\mathbb{Z}}$ is the

length-k shift. To relate this to the setting of sheaves over T, we have $W = C^0(T; \mathcal{T})$, and $\mathcal{B} = H^0(E; p^*\mathcal{T}) \subseteq C^0(E; p^*\mathcal{T}) = W^{\mathbb{Z}}$. The system so defined is linear and time-invariant. Inputs and outputs are considered via subspaces of the state space, or more generally, by *variable sharing*: constraining a function on the state space to hold a particular sequence of values over time.

Willems proposed a compositional approach to systems analysis, discussing operations he called *tearing* (separating a subsystem from a larger system), *zooming* (investigating the internal properties of a black-box system), and *linking* (connecting systems together by variable sharing). This approach has attracted recent interest from the field of applied category theory, which has formulated graphical languages for linear systems that mirror the signal flow diagrams used by engineers [BE15; Erb16]. Signal flow diagrams are instantiated as string diagrams in a monoidal category generated by the symbols used in these diagrams. The relations imposed on the generators offer a syntax that allows manipulation of diagrams while preserving equivalence of systems, composition of diagrams corresponds to gluing systems via variable sharing, and a functor to an appropriate category of behaviors implements a semantics for these diagrams. Work within this model has been able to prove some nontrivial results about controllability of systems expressed by signal flow diagrams [FSR16].

We can similarly instantiate signal flow diagrams as cellular sheaves on graphs. A signal flow diagram can be seen as a decorated graph: certain edges or vertices have symbols that represent operations on signals. These operations enforce constraints on the signals associated with edges or vertices. For discrete-time linear time invariant systems, a few operations suffice to generate all possible systems. These are shown in Figure 10.3. Each block is represented by a sheaf on a small graph. The outer vertices of this sheaf have stalks $\mathbb{R}^{\mathbb{Z}}$, representing possible discrete-time signals as measured at that vertex. The sheaf structure serves to impose certain relations between signals at these ports. The scalar multiplication block, for instance, imposes the constraint that $ax[t] = y[t]$ for all t. The time shift block imposes the constraint $x[t-1] = y[t]$ for all t.

One composes these blocks by gluing the underlying graphs together. This amounts to a colimit in the category of sheaves on all graphs. In this category, the objects are pairs $(G, \mathcal{F})$ where G is a graph and $\mathcal{F}$ a sheaf on G, and morphisms $(G, \mathcal{F}) \to (H, \mathcal{G})$ are given by a cellular morphism $f : G \to H$ and a sheaf morphism $\mathcal{F} \to f^*\mathcal{G}$. There is a forgetful functor from this category to the category of 1-dimensional regular cell complexes, and there is a free functor which is both left and right adjoint to the forgetful functor, given by assigning the zero sheaf to each graph. As a result, colimits in **Shv** have underlying graphs given by the colimit in the category of cell complexes. To build signal flow diagrams, we really only

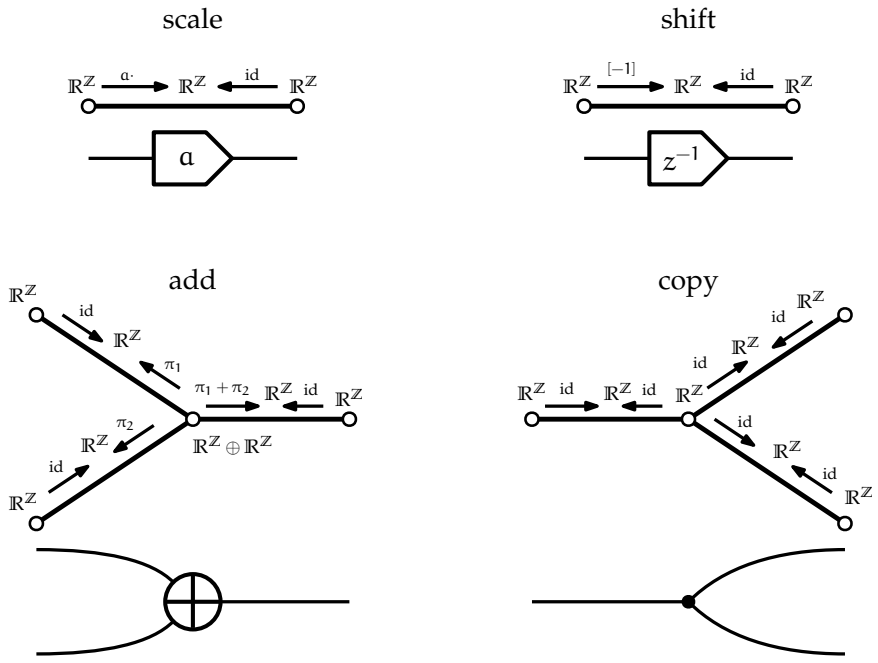


Figure 10.3: Sheaf models for the building blocks of flow diagrams for linear time-invariant systems

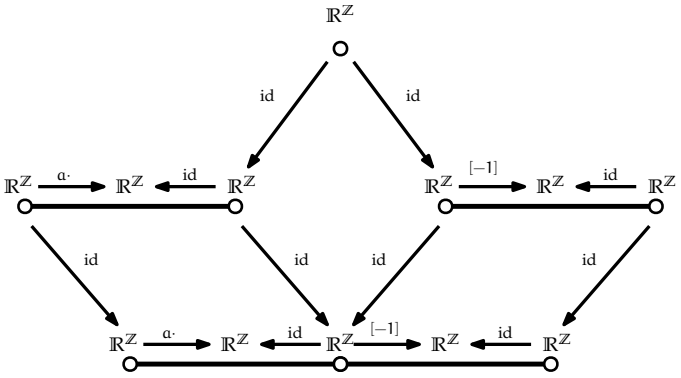


Figure 10.4: Gluing flow diagram blocks using colimits in **Shv**

need to consider diagrams in **Shv** of the form shown in Figure 10.4, for which the colimit is readily computed.

Once we have constructed our system as a pair of a graph G and a signal flow sheaf $\mathcal{S}$, we can compute its behavior by considering the space of global sections $H^0(G;\mathcal{S}) \subseteq C^0(G;\mathcal{S})$. This again gives a linear time-invariant system in the behavioral sense. However, the state space here is perhaps too large: it includes information at many internal nodes whose state is possibly unimportant to us. We would prefer to view our admissible trajectories as a subspace of a smaller space of observables. To do this, we choose a subset A of vertices of G on which the observable functions live, and take the sheaf $i^*\mathcal{S}$ on A, where $i : A \to G$ is the inclusion map. There is an induced map $i^* : H^0(G;\mathcal{S}) \to H^0(A;i^*\mathcal{S}) = C^0(A;i^*\mathcal{S})$. We let our behavior space be the image of this map within the signal space $C^0(A;i^*\mathcal{S})$.

We may also wish to construct a simplified graphical description of the system. This can be accomplished by a morphism $p : G \to H$ which collapses parts of the diagram together. The pushforward $p_*\mathcal{S}$ has the same space of global sections and hence the same behavior as the original signal flow sheaf. We may think of the graph H as representing a spatial distribution of the original system—each node represents some localized piece of the system, which is connected to others via edges.

Another way to construct a signal flow-type sheaf describing a linear system is to take a collection of linear systems Σ_i described in the format (10.1). That is, Σ_i is the system

$$\begin{aligned} x_i[t+1] &= A_i x_i[t] + B_i u_i[t] \\ y_i[t] &= C_i x_i[t] + D_i u_i[t] \end{aligned}$$

We connect these systems by attaching outputs to inputs. Suppose the state space of Σ_i is X_i, the input space is U_i, and the output space is Y_i. We take a direct sum decomposition of each $U_i = \bigoplus_j U_i^j \oplus U_i^0$ and $Y_i = \bigoplus_j Y_i^j \oplus Y_i^0$. The coupling is accomplished via linear maps $T_i^j : Y_i^j \to U_j^i$.

This collection of coupled linear systems becomes a cellular sheaf on the graph G with vertices corresponding to systems and edges between two systems when there is a nontrivial coupling (i. e., either T_i^j or T_j^i is nonzero). The stalks are $\mathcal{S}(i) = (X_i \oplus U_i)^{\mathbb{Z}}$ and $\mathcal{S}(i \sim j) = U_i^j \oplus U_j^i$. The restriction map $\mathcal{S}_{i \trianglelefteq i\sim j}$ is $\mathrm{id}_{U_i} \oplus T_i^j$. Finally, we add an additional edge e_i incident to each vertex to enforce the time evolution constraints for that system. The vertex at the opposite end will have a zero-dimensional stalk, and $\mathcal{S}(e_i) = X_i^{\mathbb{Z}}$. The restriction map $\mathcal{S}_{i \trianglelefteq e_i}$ is given by $S_{i \trianglelefteq e_i}(x[t], u[t])_{t\in\mathbb{Z}} = (x[t] - A_i x[t-1] - B_i u[t-1])_{t\in\mathbb{Z}}$. One can readily check that the sections of $\mathcal{S}$ are precisely the admissible trajectories of the coupled system.

Another way to obtain a similar sheaf is to represent each linear system as a sheaf $\mathcal{T}_i$ over the graph T in Figure 10.2. The sheaf $p_* p^* \mathcal{T}$ has a space of sections equal to the space of admissible trajectories of the linear system Σ_i. The "whiskers" of each copy of T can then be glued together in a way respecting the interconnections between systems. (More such whiskers can be added if necessary to allow for finer partitions of the input space.)

10.2.1 *Laplacians*

At this point it is natural to ask what properties the Laplacians of these sheaves associated with linear systems have. We note first that the evolution sheaf $\mathcal{E}$ has an infinite underlying graph, and the signal flow sheaf $\mathcal{S}$ has infinite-dimensional stalks, so we should expect some difficulties. We will need to restrict to signals in ℓ^2—those which decay sufficiently rapidly at infinity. Thus, the sections of $\mathcal{E}$ are those trajectories of the system with finite total energy, and we will have to replace the stalks $\mathbb{R}^{\mathbb{Z}}$ of $\mathcal{S}$ with $\ell^2(\mathbb{Z})$, at which point its sections will again be trajectories with finite total energy.

For use in applications, it will typically be necessary to truncate our spaces to a finite time horizon. Thus the graph E becomes finite and the stalks of $\mathcal{S}$ become finite-dimensional. If we are given a signal flow sheaf in the form of coupled linear systems in standard form (10.1), truncated to times t with $0 \leqslant t \leqslant T$, the space of 0-cochains can be written

$$C^0(G; \mathcal{S}) = \bigoplus_i (X_i \oplus U_i)^{T+1},$$

so we can identify a 0-cochain by its entries $(x,u)_i[t]$. The Laplacian of $\mathcal{S}$ is then given by

$$\begin{aligned}(L_{\mathcal{S}}(x,u))^x_i[t] = &\sum_{i\sim j} C_i^T(\underline{T}_i^j)^T(\underline{T}_i^j C_i x_i[t] + \underline{T}_i^j D_i u_i[t] - u_j[t]) \\ &- A_i x_i[t-1] - B_i u_i[t-1] + (I + A_i^T A_i)x_i[t] \\ &+ A_i^T B_i u_i[t] - A_i^T x_i[t+1]\end{aligned}$$

$$\begin{aligned}(L_{\mathcal{S}}(x,u))^u_i[t] = &\sum_{i\sim j} D_i^T(\underline{T}_i^j)^T(\underline{T}_i^j C_i x_i[t] + \underline{T}_i^j D_i u_i[t] - u_j[t]) \\ &+ \sum_{i\sim j}(u_i[t] - \underline{T}_j^i C_j x_j[t] - \underline{T}_j^i D_j u_j[t]) \\ &+ B_i^T A_i x_i[t] + B_i^T B_i u_i[t] - B_i^T x_i[t+1],\end{aligned}$$

where $\underline{T}_i^j$ is the extension of the map $T_i^j : Y_i^j \to U_j^i$ to $\underline{T}_i^j : Y_i \to U_j$ by zero, and all variables with $t < 0$ or $t > T$ are zero. This Laplacian is useful in the distributed model predictive control formulation described in Chapter 11.1.2.

11

DISTRIBUTED OPTIMIZATION

11.1 HOMOLOGICAL CONSTRAINTS FOR OPTIMIZATION

Consider a collection of autonomous computing agents connected in a network described by a graph G. Each agent i has access to a convex function $f_i : \mathbb{R}^n \to \mathbb{R}$. We wish to design a method by which the agents may cooperate to find the minimum of the function $f : \mathbb{R}^n \to \mathbb{R}$ given by $f(x) = \sum_i f_i(x)$, without centralized coordination.

To do this, we reformulate the problem in terms of local variables. Rather than the global problem

$$\min_{x \in \mathbb{R}^n} \sum_i f_i(x),$$

we use the equivalent problem

$$\begin{aligned} &\min_{x_i \in \mathbb{R}^n} \sum_i f_i(x_i) \\ &\quad \text{s.t. } x_i = x_j \text{ for all } i \sim j. \end{aligned} \tag{11.1}$$

We can think of these distributed variables as representing a cochain $x \in C^0(G; \underline{\mathbb{R}^n})$. The equality constraints then amount to the requirement that $x \in H^0(G; \underline{\mathbb{R}^n})$.

This reformulation turns the problem into a naturally decentralized problem. Distributed optimization algorithms have been studied extensively, and typically involve a combination of a local optimization algorithm acting on each f_i—like gradient descent—and a neighborhood-based consensus process—like graph diffusion. When appropriately combined, the first process moves toward a minimizer for the problem, while the second process ensures that the homological constraint is satisfied.

While the initial motivation comes from a constraint $x \in H^0(G; \underline{\mathbb{R}^n})$, we may just as easily use any other sheaf on G. We call this type of problem *homological programming*, by analogy with the terminology for other sorts of optimization problems. This naturally extends to optimization problems defined on higher-degree cochains. Given an objective $f(x) = \sum_{\dim \sigma = d} f_\sigma(x_\sigma)$, we can optimize

subject to the constraint $x \in \mathcal{H}^d(X;\mathcal{F})$ or $x \in Z^k(X;\mathcal{F})$, or even $x \perp B^k(X;\mathcal{F})$. These constraints can be enforced by the Hodge, up-, or down-Laplacians, respectively.

The framework of homological programming subsumes that of distributed optimization with a single state space by using the constant sheaf, but it also includes distributed formulations of other natural problems. It should be noted that the question of algorithms for optimization problems with distributed linear constraints has been considered previously, in [Nas+17; Hua+17]. These papers considered distributed learning problems where local learned parameters were required to satisfy certain linear constraints between sets of nodes.

Here are a couple of sketches to illustrate potential uses of sheaf homological programs.

11.1.1 *Signal recovery*

Consider a set A of sensors observing a function $\mathcal{D} \to \mathbb{V}$ defined on some domain $\mathcal{D}$. Each sensor $a \in A$ observes f restricted to some subset U_a of $\mathcal{D}$, and assume that the sets U_a cover $\mathcal{D}$. The nerve of this cover is a simplicial complex $\mathcal{N}(A)$ with vertices corresponding to sensors. This nerve carries a natural cellular sheaf $\mathcal{F}$. The stalk over a cell σ corresponding to an intersection of sets $\bigcap_A U_a$ is simply the set of functions $\bigcap_I U_i \to \mathbb{V}$. Restriction maps are simply restriction of functions. Global sections of $\mathcal{F}$ are precisely functions $\mathcal{D} \to \mathbb{V}$.

We now assume that the sensors have the ability to communicate with their neighbors in the 1-skeleton of the nerve. Assume each sensor a can observe $x_a = P_a f|_{U_a} + \epsilon_a$, where P_a is a linear function representing some sensing modality, and ϵ_a is an error term. (The sensing map P_a might simply be the identity, or it might be something more complicated like a sampling or a band-limited Fourier transform.) The problem of recovering f from these measurements can be cast as a regularized optimization problem:

$$\min_f \sum_a \|P_a f|_{U_a} - x_a\|^2 + \alpha R(f)$$

where $R(f)$ is some regularization term controlling properties of the recovered solution. If we assume $R(f)$ is local, consisting of a sum of terms $R_a(f|_{U_a})$, we can write an equivalent homological program

$$\begin{aligned} \min_{y \in C^0(\mathcal{N}(A);\mathcal{F})} & \sum_a \|P_a y_a - x_a\|^2 + \alpha R_a(y_a) \\ \text{s.t. } & y \in H^0(\mathcal{N}(A);\mathcal{F}) \end{aligned}$$

For a more concrete instantiation of this problem, consider the case where A and $\mathcal{D}$ are the set of vertices of a graph G, and U_a for a vertex a is the set of neighbors of a. That is, sensors are located at the vertices of a graph, and can observe (with noise) the values of f at all vertices in their immediate neighborhood. We take $P_a = \mathrm{id}$, and assume that the observed function is smooth, so that our regularization term is $R(f) = \langle f, L_G f \rangle = \sum_{v \in V(G)} \frac{1}{2} \sum_{u \sim v} (f(v) - f(u))^2 = \sum_v R_v(f|_{U_v})$. Then the distributed optimization problem is

$$\begin{aligned} \min_{y \in C^0(G;\mathcal{F})} & \sum_v \|y_v - x_v\|^2 + \frac{\alpha}{2} \sum_{u \sim v} (y_v(v) - y_v(u))^2 \\ \text{s.t. } & y \in H^0(G;\mathcal{F}) \end{aligned}$$

11.1.2 *Model predictive control*

Optimal control theory considers the problem of controlling a dynamical system to a particular state at a minimal cost. For a linear system with no constraints on inputs, these problems can be solved explicitly. Feedback controllers can also be designed to stabilize the system around a particular operating point. However, when the system model is subject to inaccuracies or has high-order behavior, these methods may not suffice to control the system. In these situations, a more robust approach known as *model predictive control* is often used.

In model predictive control, the input at each time step is selected based on a finite-horizon optimization problem based on information at the current time. That is, given the observed state of the system at time t, compute inputs for times $t+1, \ldots, t+T$ that minimize an objective function designed to steer the output to a desired state over this time horizon. Formally, one considers an optimization problem like

$$\min_u \sum_{t=1}^{T} C_x(x[t]) + C_u(u)$$

with the implicit constraint that the state variables $x[t]$ depend on the input variables $u[t]$ according to the linear model for the system. At each time step only the first input from the calculated plan is applied to the system, and then the optimization problem is solved again at the next step with the new measurements.

Given a collection of coupled systems with local controllers that can communicate, we might want a distributed method of solving the model predictive control problem. This requires a distributed way of computing the model relationship between inputs and states. The signal flow sheaf of Chapter 10.2 offers just such a

method of computation. Under the assumption that the cost functions C_x and C_u decompose with respect to local states and inputs, the MPC problem becomes a homological program:

$$\begin{aligned} \min_{(x,u)\in C^0(G;\mathcal{S})} & \sum_i \left(\sum_{t=1}^{T} C_x^i(x_i[t]) + C_u^i(u_i) \right) \\ \text{s.t. } & (x,u) \in H^0(G;\mathcal{S}) \\ & x[0] = x_0. \end{aligned} \tag{11.2}$$

There is an extra linear constraint here, which can be treated in various ways depending on the distributed algorithm we use. Perhaps the simplest is to fix $x[0]$ in the dynamics and use the harmonic extension Laplacian dynamics.

11.2 DISTRIBUTED ALGORITHMS FOR HOMOLOGICAL PROGRAMMING

The sheaf Laplacian implements the constraints for homological programs in the same way that the graph Laplacian implements the local constancy constraint. This allows us to adapt algorithms for distributed optimization to homological programming. One such approach, proposed by Wang and Elia [WE10; WE11], constructs a distributed dynamical system that converges to the optimizer of the problem. Given a problem of the form

$$\begin{aligned} \min_{x_i\in\mathbb{R}^n} & \sum_i f_i(x_i) \\ \text{s.t. } & Lx = 0 \end{aligned} \tag{11.3}$$

construct an augmented Lagrangian

$$\mathcal{L}(x,z) = \sum_i f_i(x_i) + \langle x, Lx\rangle + \langle z, Lx\rangle.$$

This function is convex in x and linear in z; saddle points of the function are primal-dual optimal pairs for the optimization problem. The saddle-point dynamics

$$\begin{aligned} \dot{x} &= -\nabla_x \mathcal{L}(x,z) = -\sum_i \nabla f_i(x_i) - 2Lx - Lz \\ \dot{z} &= \nabla_z \mathcal{L}(x,z) = Lx \end{aligned}$$

are gradient descent in the primal variable and gradient ascent in the dual variable. These dynamics are local, since the derivative of x_i depends only on ∇f_i, $(Lx)_i$ and $(Lz)_i$. A result about saddle-point dynamics due to Cherukuri, Gharesifard, and

Cortés allows us to show that this system indeed converges globally to optimal primal-dual pairs.

Theorem 11.2.1 (Corollary 4.5, [CGC17]). *Let $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ be once differentiable, with the function $\mathcal{L}(\cdot, z)$ convex for all z and the function $\mathcal{L}(x, \cdot)$ linear. Further suppose that for each saddle point (x_*, z_*) of $\mathcal{L}$, if $\mathcal{L}(x, z_*) = \mathcal{L}(x_*, z_*)$ then (x, z_*) is a saddle point of $\mathcal{L}$. Then the set of saddle points of $\mathcal{L}$ is globally asymptotically stable under the saddle-point dynamics, with trajectories converging to a point.*

The saddle points of $\mathcal{L}$ are precisely the optimal primal-dual pairs for the augmented problem. Since the objective function is convex, the conditions for this theorem hold.

Of course, there is no formal difference between the system 11.3 and the system

$$\begin{aligned} &\min_{x \in C^0(G;\mathcal{F})} \sum_v f_v(x_v) \\ &\quad \text{s.t. } L_{\mathcal{F}} x = 0, \end{aligned} \tag{11.4}$$

so this method immediately generalizes to give a distributed algorithm for homological programming. All one has to do is replace the graph Laplacian with the sheaf Laplacian in the augmented Lagrangian.

BIBLIOGRAPHY

[AC88] N. Alon and F. R. K. Chung. "Explicit Construction of Linear Sized Tolerant Networks." In: *Discrete Mathematics* 72.1 (Dec. 1988), pp. 15–19.

[AKR06] T. Antal, P. L. Krapivsky, and S. Redner. "Social Balance on Networks: The Dynamics of Friendship and Enmity." In: *Physica D: Nonlinear Phenomena*. Dynamics on Complex Networks and Applications 224.1 (Dec. 2006), pp. 130–136.

[ABRK19] Fouzul Atik, R. B. Bapat, and M. Rajesh Kannan. "Resistance Matrices of Graphs with Matrix Weights." In: *Linear Algebra and its Applications* 571 (June 2019), pp. 41–57.

[BE15] John C. Baez and Jason Erbele. "Categories in Control." In: *arXiv:1405.6881* (May 2015). arXiv: 1405.6881.

[BSS13] Afonso S. Bandeira, Amit Singer, and Daniel A. Spielman. "A Cheeger Inequality for the Graph Connection Laplacian." In: *SIAM Journal on Matrix Analysis and Applications* 34.4 (2013), pp. 1611–1630.

[BGd08] Onureena Banerjee, Laurent El Ghaoui, and Alexandre d'Aspremont. "Model Selection through Sparse Maximum Likelihood Estimation for Multivariate Gaussian or Binary Data." In: *Journal of Machine Learning Research* 9.Mar (2008), pp. 485–516.

[BSB09] S. Barbarossa, G. Scutari, and T. Battisti. "Distributed Signal Subspace Projection Algorithms with Maximum Convergence Rate for Sensor Networks with Topological Constraints." In: *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. Taipei, Taiwan: IEEE, Apr. 2009, pp. 2893–2896.

[Bär15] Manuel Bärenz. *MathOverflow Answer: Are Dagger Categories Truly Evil?* Dec. 2015.

[BSS12] Joshua Batson, Daniel A. Spielman, and Nikhil Srivastava. "Twice-Ramanujan Sparsifiers." In: *SIAM Journal on Computing* 41.6 (2012), pp. 1704–1721.

[BN03] Mikhail Belkin and Partha Niyogi. "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation." In: *Neural Computation* 15 (2003), pp. 1373–1396.

[BL06] Yonatan Bilu and Nathan Linial. "Lifts, Discrepancy and Nearly Optimal Spectral Gap." In: *Combinatorica* 26.5 (Oct. 2006), pp. 495–519.

[Bjö84] A. Björner. "Posets, Regular CW Complexes and Bruhat Order." In: *European Journal of Combinatorics* 5.1 (1984), pp. 7–16.

[Bom+05] Erik G. Boman, Doron Chen, Ojas Parekh, and Sivan Toledo. "On Factor Width and Symmetric H-Matrices." In: *Linear Algebra and its Applications* 405.1 (2005), pp. 239–248.

[Boy+05] S. Boyd, A. Ghosh, B. Prabhakar, and D. Shah. "Gossip Algorithms: Design, Analysis and Applications." In: *Proceedings IEEE 24th Annual Joint Conference of the IEEE Computer and Communications Societies.* Vol. 3. Mar. 2005, 1653–1664 vol. 3.

[BL92] J. Brüning and M. Lesch. "Hilbert Complexes." In: *Journal of Functional Analysis* 108.1 (Aug. 1992), pp. 88–132.

[Bul19] Francesco Bullo. *Lectures on Network Systems*. Version 1.3. Kindle Direct Publishing, 2019.

[CSHS16] Marcel K. de Carli Silva, Nicholas J. A. Harvey, and Christiane M. Sato. "Sparse Sums of Positive Semidefinite Matrices." In: *ACM Transactions on Algorithms* 12.1 (2016), 9:1–9:17.

[CH56] Dorwin Cartwright and Frank Harary. "Structural Balance: A Generalization of Heider's Theory." In: *Psychological Review* 63 (1956), pp. 277–293.

[Cas+11] Peter G. Casazza, Matthew Fickus, Dustin G. Mixon, Yang Wang, and Zhengfang Zhou. "Constructing Tight Fusion Frames." In: *Applied and Computational Harmonic Analysis* 30.2 (Mar. 2011), pp. 175–187.

[CF11] Francesca Ceragioli and Paolo Frasca. "Continuous-Time Discontinuous Equations in Bounded Confidence Opinion Dynamics." In: *IFAC Proceedings Volumes*. 18th IFAC World Congress 44.1 (Jan. 2011), pp. 1986–1990.

[CGC17] A. Cherukuri, B. Gharesifard, and J. Cortés. "Saddle-Point Dynamics: Conditions for Asymptotic Stability of Saddle Points." In: *SIAM Journal on Control and Optimization* 55.1 (Jan. 2017), pp. 486–511.

[CK15] Fan Chung and Mark Kempton. "A Local Clustering Algorithm for Connection Graphs." In: *Internet Mathematics* 11.4-5 (Sept. 2015), pp. 333–351.

[CZ12] Fan Chung and Wenbo Zhao. "Ranking and Sparsifying a Connection Graph." In: *International Workshop on Algorithms and Models for the Web-Graph*. Lecture Notes in Computer Science. Springer, 2012.

[CL06] Ronald R. Coifman and Stéphane Lafon. "Diffusion Maps." In: *Applied and Computational Harmonic Analysis* 21.1 (July 2006), pp. 5–30.

[Cur15] Justin Michael Curry. "Topological Data Analysis and Cosheaves." In: *Japan Journal of Industrial and Applied Mathematics* 32.2 (July 2015), pp. 333–371.

[Cur14] Justin Curry. "Sheaves, Cosheaves, and Applications." PhD thesis. University of Pennsylvania, 2014.

[DeG74] Morris H. DeGroot. "Reaching a Consensus." In: *Journal of the American Statistical Association* 69 (Mar. 1974), pp. 118–121.

[Dit01] Jan Christian Dittmer. "Consensus Formation under Bounded Confidence." In: *Nonlinear Analysis: Theory, Methods & Applications*. Proceedings of the Third World Congress of Nonlinear Analysts 47.7 (Aug. 2001), pp. 4615–4621.

[Dod76] Jozef Dodziuk. "Finite-Difference Approach to the Hodge Theory of Harmonic Forms." In: *American Journal of Mathematics* 98.1 (1976), p. 79.

[Dod84] Jozef Dodziuk. "Difference Equations, Isoperimetric Inequality and Transience of Certain Random Walks." en. In: *Transactions of the American Mathematical Society* 284.2 (1984).

[Don+15] Xiaowen Dong, Dorina Thanou, Pascal Frossard, and Pierre Vandergheynst. "Laplacian Matrix Learning for Smooth Graph Signal Representation." In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. South Brisbane, Queensland, Australia: IEEE, Apr. 2015, pp. 3736–3740.

[Eck45] Beno Eckmann. "Harmonische Funktionen Und Randwertaufgaben in Einem Komplex." In: *Commentarii Mathematici Helvetici* 17.1 (1945), pp. 240–255.

[EPO17] H. E. Egilmez, E. Pavez, and A. Ortega. "Graph Learning from Data under Laplacian and Structural Constraints." In: *IEEE Journal of Selected Topics in Signal Processing* 11.6 (Sept. 2017), pp. 825–841.

[EKW15] Noureddine El Karoui and Hau-tieng Wu. "Graph Connection Laplacian and Random Matrices with Random Blocks." In: *Information and Inference: A Journal of the IMA* 4.1 (Mar. 2015), pp. 1–44.

[Erb16] Jason Michael Erbele. "Categories in Control: Applied PROPs." In: *arXiv:1611.07591 [math]* (Nov. 2016). arXiv: 1611.07591 [math].

[ET15] Brent Everitt and Paul Turner. "Cellular Cohomology of Posets with Local Coefficients." In: *Journal of Algebra* 439 (Oct. 2015), p. 20.

[FF17] Nicolas Folinsbee and Joel Friedman. "Sheaves and Duality in the Two-Vertex Graph Riemann-Roch Theorem." In: *arXiv:1712.08695 [math]* (Dec. 2017). arXiv: 1712.08695 [math].

[FSR16] Brendan Fong, Paweł Sobociński, and Paolo Rapisarda. "A Categorical Approach to Open and Interconnected Dynamical Systems." In: *2016 31st Annual ACM/IEEE Symposium on Logic in Computer Science (LICS)*. July 2016, pp. 1–10.

[FHT08] Jerome Friedman, Trevor Hastie, and Robert Tibshirani. "Sparse Inverse Covariance Estimation with the Graphical Lasso." In: *Biostatistics* 9.3 (July 2008), pp. 432–441.

[Fri15] Joel Friedman. "Sheaves on Graphs, Their Homological Invariants, and a Proof of the Hanna Neumann Conjecture." In: *Memoirs of the American Mathematical Society* 233.1100 (Jan. 2015).

[Gao19] Tingran Gao. "The Diffusion Geometry of Fibre Bundles: Horizontal Diffusion Maps." In: *Applied and Computational Harmonic Analysis* (Aug. 2019).

[GBM19] Tingran Gao, Jacek Brodzki, and Sayan Mukherjee. "The Geometry of Synchronization Problems and Learning Group Actions." In: *Discrete & Computational Geometry* (May 2019).

[Gar84] L. Terrell Gardner. "Shorter Notes: An Elementary Proof of the Russo-Dye Theorem." In: *Proceedings of the American Mathematical Society* 90.1 (1984), pp. 171–171.

[GS10] Federica Garin and Luca Schenato. "A Survey on Distributed Estimation and Control Applications Using Linear Consensus Algorithms." In: *Networked Control Systems*. Ed. by Alberto Bemporad, Maurice Heemels, and Mikael Johansson. Lecture Notes in Control and Information Sciences. London: Springer, 2010, pp. 75–107.

[GM03] Sergei I. Gelfand and Yuri I. Manin. *Methods of Homological Algebra*. Second edition. Springer Monographs in Mathematics. Berlin: Springer-Verlag, 2003.

[Gro09] Mikhail Gromov. "Singularities, Expanders and Topology of Maps. Part 1: Homology versus Volume in the Spaces of Cycles." In: *Geometric and Functional Analysis* 19.3 (Dec. 2009), pp. 743–841.

[Gro10] Mikhail Gromov. "Singularities, Expanders and Topology of Maps. Part 2: From Combinatorics to Topology via Algebraic Isoperimetry." In: *Geometric and Functional Analysis* 20 (2010), pp. 416–526.

[GT87] Jonathan L. Gross and Thomas W. Tucker. *Topological Graph Theory*. Wiley-Interscience Series in Discrete Mathematics and Optimization. New York: John Wiley & Sons, 1987.

[Har53] Frank Harary. "On the Notion of Balance of a Signed Graph." In: *The Michigan Mathematical Journal* 2.2 (1953), pp. 143–146.

[HK02] Rainer Hegselmann and Ulrich Krause. "Opinion Dynamics and Bounded Confidence Models, Analysis and Simulation." In: *Journal of Artificial Societies and Social Simulation* 5.3 (June 2002).

[HK19] Chris Heunen and Matti Karvonen. "Limits in Dagger Categories." In: *Theory and Applications of Categories* 34.18 (2019), pp. 468–513.

[HJ13] Danijela Horak and Jürgen Jost. "Spectra of Combinatorial Laplace Operators on Simplicial Complexes." In: *Advances in Mathematics* 244 (2013), pp. 303–336.

[Hor11] Steve Horvath. *Weighted Network Analysis: Applications in Genomics and Systems Biology*. New York: Springer-Verlag, 2011.

[Hua+17] F. Hua, R. Nassif, C. Richard, H. Wang, and J. Huang. "Penalty-Based Multitask Distributed Adaptation over Networks with Constraints." In: *2017 51st Asilomar Conference on Signals, Systems, and Computers*. Oct. 2017, pp. 908–912.

[HLW19] Vahan Huroyan, Gilad Lerman, and Hau-Tieng Wu. "Solving Jigsaw Puzzles by the Graph Connection Laplacian." In: *arXiv:1811.03188 [cs, math]* (Nov. 2019). arXiv: 1811.03188 [cs, math].

[JLM03] Ali Jadbabaie, Jie Lin, and A Stephen Morse. "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules." In: *IEEE Transactions on Automatic Control* 48.6 (2003), p. 16.

[Jew62] William S. Jewell. "Optimal Flow through Networks with Gains." In: *Operations Research* 10.4 (1962), pp. 476–499.

[Jia+11] Xiaoye Jiang, Lek-Heng Lim, Yuan Yao, and Yinyu Ye. "Statistical Ranking and Combinatorial Hodge Theory." In: *Mathematical Programming* 127.1 (Mar. 2011), pp. 203–244.

[JS05] Charles R. Johnson and Ronald L. Smith. "Closure Properties." In: *The Schur Complement and Its Applications*. Numerical Methods and Algorithms 4. Springer, 2005, pp. 111–136.

[Kac66] Mark Kac. "Can One Hear the Shape of a Drum?" In: *The American Mathematical Monthly* 73.4 (Apr. 1966), pp. 1–23.

[Kal16] Vassilis Kalofolias. "How to Learn a Graph from Smooth Signals." In: *Journal of Machine Learning Research (JMLR)* (2016).

[Kal15] Kaushik Kalyanaraman. "Hodge Laplacians on Simplicial Meshes and Graphs." PhD thesis. Urbana, IL: University of Illinois at Urbana-Champaign, 2015.

[KS90] Masaki Kashiwara and Pierre Schapira. *Sheaves on Manifolds*. Grundlehren Der Mathematischen Wissenschaften 292. Springer-Verlag Berlin Heidelberg, 1990.

[Kel67] Al Kelley. "Stability of the Center-Stable Manifold." In: *Journal of Mathematical Analysis and Applications* 18.2 (May 1967), pp. 336–344.

[Kem15] Mark C. Kempton. "High Dimensional Spectral Graph Theory and Non-Backtracking Random Walks on Graphs." PhD thesis. UC San Diego, 2015.

[Kho10] Subhash Khot. "On the Unique Games Conjecture." In: *2010 IEEE 25th Annual Conference on Computational Complexity*. June 2010, pp. 99–121.

[KMS17] Subhash Khot, Dor Minzer, and Muli Safra. "On Independent Sets, 2-to-2 Games, and Grassmann Graphs." In: *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*. STOC 2017. Montreal, Canada: Association for Computing Machinery, June 2017, pp. 576–589.

[KMS18] Subhash Khot, Dor Minzer, and Muli Safra. *Pseudorandom Sets in Grassmann Graph Have Near-Perfect Expansion*. Tech. rep. 006. 2018.

[Kiv+14] Mikko Kivelä, Alexandre Arenas, Marc Barthelemy, James P. Gleeson, Yamir Moreno, and Mason A. Porter. "Multilayer Networks." In: *Journal of Complex Networks* 2.3 (Sept. 2014), pp. 203–271. arXiv: 1309.7233.

[KB67] A. N. Kolmogorov and Y. M. Barzdin. "On the Realization of Nets in 3-Dimensional Space." In: *Problems in Cybernetics* 8 (1967), pp. 261–268.

[Kyn+16] Rasmus Kyng, Yin Tat Lee, Richard Peng, Sushant Sachdeva, and Daniel A. Spielman. "Sparsified Cholesky and Multigrid Solvers for Connection Laplacians." In: *Proceedings of the Forty-Eighth Annual ACM Symposium on Theory of Computing*. STOC '16. Cambridge, MA, USA: Association for Computing Machinery, June 2016, pp. 842–850.

[LT10] Brenden Lake and Joshua Tenenbaum. "Discovering Structure by Learning Sparse Graphs." In: *Proceedings of the Annual Meeting of the Cognitive Science Society* 32.32 (2010).

[LBS19] Paolo Di Lorenzo, Sergio Barbarossa, and Stefania Sardellitti. "Distributed Signal Recovery Based on In-Network Subspace Projections." In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Brighton, United Kingdom: IEEE, May 2019, pp. 5242–5246.

[LPS88] A. Lubotzky, R. Phillips, and P. Sarnak. "Ramanujan Graphs." In: *Combinatorica* 8.3 (Sept. 1988), pp. 261–277.

[LW69] Albert T. Lundell and Stephen Weingram. *The Topology of CW Complexes*. New York, NY: Van Nostrand Reinhold, 1969.

[MJW06] Dmitry M Malioutov, Jason K Johnson, and Alan S Willsky. "Walk-Sums and Belief Propagation in Gaussian Graphical Models." In: *Journal of Machine Learning Research* 7 (2006), pp. 2031–2064.

[Man08] Tatiana Mantuano. "Discretization of Riemannian Manifolds Applied to the Hodge Laplacian." In: *American Journal of Mathematics* 130.6 (2008), pp. 1477–1508.

[MSS15] Adam W. Marcus, Daniel A. Spielman, and Nikhil Srivastava. "Interlacing Families I: Bipartite Ramanujan Graphs of All Degrees." In: *Annals of Mathematics* 182.1 (2015), pp. 307–325.

[MSS18] Adam W. Marcus, Daniel A. Spielman, and Nikhil. Srivastava. "Interlacing Families IV: Bipartite Ramanujan Graphs of All Sizes." In: *SIAM Journal on Computing* 47.6 (Jan. 2018), pp. 2488–2509.

[Mar88] G.A. Margulis. "Explicit Group-Theoretical Constructions of Combinatorial Schemes and Their Application to the Design of Expanders and Concentrators." In: *Problems in Information Transmission* 24.1 (1988), pp. 39–46.

[McL03] Colin McLarty. *The Rising Sea: Grothendieck on Simplicity and Generality I*. 2003.

[NVS19] R. Nassif, S. Vlaski, and A. H. Sayed. "Distributed Inference over Networks under Subspace Constraints." In: *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. May 2019, pp. 5232–5236.

[Nas+17] R. Nassif, C. Richard, A. Ferrari, and A. H. Sayed. "Diffusion LMS for Multitask Problems with Local Linear Equality Constraints." In: *IEEE Transactions on Signal Processing* 65.19 (Oct. 2017), pp. 4979–4993.

[Nic07] Liviu I Nicolaescu. *Lectures on the Geometry of Manifolds*. Second edition. World Scientific, Sept. 2007.

[Nil91] A. Nilli. "On the Second Eigenvalue of a Graph." In: *Discrete Mathematics* 91.2 (Aug. 1991), pp. 207–210.

[OM04] R. Olfati-Saber and R.M. Murray. "Consensus Problems in Networks of Agents with Switching Topology and Time-Delays." In: *IEEE Transactions on Automatic Control* 49.9 (Sept. 2004), pp. 1520–1533.

[OT11] Alex Olshevsky and John N. Tsitsiklis. "Convergence Speed in Distributed Consensus and Averaging." In: *SIAM Review* 53.4 (Jan. 2011), pp. 747–772.

[Ort+18] A. Ortega, P. Frossard, J. Kovačević, J. M. F. Moura, and P. Vandergheynst. "Graph Signal Processing: Overview, Challenges, and Applications." In: *Proceedings of the IEEE* 106.5 (May 2018), pp. 808–828.

[OPW19] Braxton Osting, Sourabh Palande, and Bei Wang. "Spectral Sparsification of Simplicial Complexes for Clustering and Label Propagation." In: *arXiv:1708.08436 [cs]* (Feb. 2019). arXiv: 1708.08436 [cs].

[Par13] Ori Parzanchevski. "High Dimensional Expanders." PhD thesis. Hebrew University of Jerusalem, 2013.

[Pin73] Mark S. Pinsker. "On the Complexity of a Concentrator." In: *7th International Teletraffic Conference*. 1973.

[RWW18] Shiquan Ren, Chengyuan Wu, and Jie Wu. "Weighted Persistent Homology." In: *arXiv:1708.06722 [math]* (Apr. 2018). arXiv: 1708.06722 [math].

[RV07] Mark Rudelson and Roman Vershynin. "Sampling from Large Matrices: An Approach through Geometric Functional Analysis." In: *Journal of the ACM* 54.4 (July 2007), 21–es.

[RH05] Håvard Rue and Leonard Held. *Gaussian Markov Random Fields: Theory and Applications*. Monographs on Statistics and Applied Probability 104. CRC Press, 2005.

[She85] Allen Shepard. "A Cellular Description of the Derived Category of a Stratified Space." PhD thesis. Brown University, May 1985.

[SMP16] Vin de Silva, Elizabeth Munch, and Amit Patel. "Categorified Reeb Graphs." In: *Discrete & Computational Geometry* 55.4 (June 2016), pp. 854–906.

[Sin11] A. Singer. "Angular Synchronization by Eigenvectors and Semidefinite Programming." In: *Applied and Computational Harmonic Analysis* 30.1 (Jan. 2011), pp. 20–36.

[SW12] Amit Singer and Hau-Tieng Wu. "Vector Diffusion Maps and the Connection Laplacian." In: *Communications in Pure and Applied Mathematics* 65.8 (2012).

[Son98] Eduardo D. Sontag. *Mathematical Control Theory: Deterministic Finite Dimensional Systems*. Second. Texts in Applied Mathematics. New York: Springer-Verlag, 1998.

[SS11] Daniel A. Spielman and Nikhil. Srivastava. "Graph Sparsification by Effective Resistances." In: *SIAM Journal on Computing* 40.6 (Jan. 2011), pp. 1913–1926.

[ST11] Daniel Spielman and Shang-Hua Teng. "Spectral Sparsification of Graphs." In: *SIAM Journal on Computing* 40.4 (2011), pp. 981–1025.

[Ste13] John Steenbergen. "Towards a Spectral Theory for Simplicial Complexes." PhD thesis. Duke University, 2013.

[Suno8] Toshikazu Sunada. "Discrete Geometric Analysis." In: *Proceedings of Symposia in Pure Mathematics* (Jan. 2008).

[Tay68] Michael Taylor. "Towards a Mathematical Theory of Influence and Attitude Change." In: *Human Relations* 21.2 (May 1968), pp. 121–139.

[Tri+18] Minh Hoang Trinh, Chuong Van Nguyen, Young-Hun Lim, and Hyo-Sung Ahn. "Matrix-Weighted Consensus and Its Applications." In: *Automatica* 89 (Mar. 2018), pp. 415–419.

[Tun16] S. Emre Tuna. "Synchronization under Matrix-Weighted Laplacian." In: *Automatica* 73 (Nov. 2016), pp. 76–81.

[WE10] J. Wang and N. Elia. "Control Approach to Distributed Optimization." In: *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. Sept. 2010, pp. 557–561.

[WE11] J. Wang and N. Elia. "A Control Perspective for Centralized and Distributed Convex Optimization." In: *2011 50th IEEE Conference on Decision and Control and European Control Conference*. Dec. 2011, pp. 3800–3805.

[Wil07] J. C. Willems. "The Behavioral Approach to Open and Interconnected Systems." In: *IEEE Control Systems Magazine* 27.6 (Dec. 2007), pp. 46–99.

[Wil86] Jan C. Willems. "From Time Series to Linear System—Part I. Finite Dimensional Linear Time Invariant Systems." In: *Automatica* 22.5 (Sept. 1986), pp. 561–580.

[Wu+19] Chengyuan Wu, Shiquan Ren, Jie Wu, and Kelin Xia. "Weighted Fundamental Group." In: *arXiv:1808.06184 [math]* (Apr. 2019). arXiv: 1808.06184 [math].

[Wu+20] Chengyuan Wu, Shiquan Ren, Jie Wu, and Kelin Xia. "Discrete Morse Theory for Weighted Simplicial Complexes." In: *Topology and its Applications* 270 (Feb. 2020), p. 107038.

[XB04] Lin Xiao and Stephen Boyd. "Fast Linear Iterations for Distributed Averaging." In: *Systems & Control Letters* 53.1 (Sept. 2004), pp. 65–78.

[Xu+17] Qianqian Xu, Jiechao Xiong, Xi Chen, Qingming Huang, and Yuan Yao. "HodgeRank with Information Maximization for Crowdsourced Pairwise Ranking Aggregation." In: *arXiv:1711.05957 [stat]* (Nov. 2017). arXiv: 1711.05957 [stat].

[Zas89] Thomas Zaslavsky. "Biased Graphs. I. Bias, Balance, and Gains." In: *Journal of Combinatorial Theory, Series B* 47.1 (Aug. 1989), pp. 32–52.

[Zas91] Thomas Zaslavsky. "Biased Graphs. II. The Three Matroids." In: *Journal of Combinatorial Theory, Series B* 51.1 (Jan. 1991), pp. 46–72.

[Zas02] Thomas Zaslavsky. "Glossary of Signed and Gain Graphs." In: *The Electronic Journal of Combinatorics* (2002), DS9: Sep 18–1998.

[Zas18] Thomas Zaslavsky. "A Mathematical Bibliography of Signed and Gain Graphs and Allied Areas." In: *The Electronic Journal of Combinatorics* (2018), DS8: Dec 21–2018.

[ZKC14] Wenbo Zhao, Mark Kempton, and Fan Chung. "Ranking and Sparsifying a Connection Graph." In: *Internet Mathematics* 10.1 (Apr. 2014), p. 1553.