FOUNDATIONS AND APPLICATIONS OF THE LAPLACE-BELTRAMI OPERATOR

Layla Jarrahy **Christian Lentz** **Vinith Yedidi**

ABSTRACT

The Laplace-Beltrami operator is a generalization of the standard Laplacian to functions defined over manifolds. In this review, we formulate the definition of the Laplace-Beltrami operator starting only with a definition of a topological manifold. We emphasize the connection to discrete methods of dimensionality reduction and motivate Laplace-Beltrami as a generalization of the well-known Laplacian Eigenmap (LE) and Locally Linear Embedding (LLE). Our discussion emphasizes key properties of the operator, and investigates applications including heat flow on a manifold, geometry processing, and Laplacian autoencoders. The code to generate our examples can be accessed at [27].

***K*eywords** Laplace-Beltrami operator, partial differential equations, manifold learning, dimensionality reduction, geometric machine learning, numerical algorithms

1 Introduction

Manifold learning refers to a class of machine learning approaches that operate given the underlying assumption that data lie on or near an embedded manifold [34, 25, 33]. This assumption is often referred to as the *manifold hypothesis*. It is common to refer to the dimension of the embedded manifold as the *intrinsic dimensionality* of the data. The notion of intrinsic dimensionality has been useful in the analysis of language models, for example [1].

The manifold learning paradigm is motivated by a common task in data analysis: given high-dimensional data, how can one construct lower dimensional representations which faithfully represent the original data? In the case that the intrinsic dimensionality of the data is much less than that of the ambient Euclidean space, the construction of such representations is referred to as *dimensionality reduction* (DR). Although not all methods of DR are manifold based, our analysis will exclusively examine such algorithms. These methods facilitate data analysis by reducing storage space and processing time. Furthermore, as the dimensionality of data increases, the performance of machine learning algorithms degrades [47, 54], and the complexity of learned models increases [29, 56], further motivating the need for DR. Although formalization of the manifold hypothesis is still an active area of research, certain works testing the validity of the hypothesis have provided evidence that such methods are reliable in practice [20, 34].

The development and application of DR algorithms has emerged as a particularly active area of research in recent years—especially following the seminal works introducing the *ISOMAP* [50] and *Locally Linear Embedding (LLE)* [44] algorithms. These methods emerged from a need to generalize the classical methods of Principal Components Analysis (PCA) [46, 18] and Multidimensional Scaling (MDS) [30, 24, 13] in order to analyze high-dimensional data with nonlinear manifold structure in a similar manner. A number of methods of so-called *nonlinear dimensionality reduction* (NLDR) techniques have been introduced in recent decades [45, 50, 44, 4, 53, 15]. There is empirical evidence that, for data with sufficiently high dimensionality, machine learning methods which use PCA as a pre-processing step perform better when implemented appropriately [40]. The choice of *linear dimensionality reduction* (LDR) verses NLDR method is data dependent, although certain themes hold in each category. For example, the authors of [3] perform three case studies and conclude that (1) NLDR methods often outperformed LDR methods in this setting and (2) NLDR methods excelled at uncovering local manifold structure.

In the present work, we are particularly interested in Laplacian methods including LLE [44] and the *Laplacian Eigenmap* (LE) [4]. Our primary focus will be to study a generalization of these methods to a continuous setting. We will see in Section 3 that the LLE and LE algorithms both aim to construct an orthogonal basis of vectors satisfying certain constraints. In practice, this is achieved by solving a (possibly generalized) eigenvalue problem. In fact, the *Laplace-Beltrami operator* extends this notion to a continuous regime [32, 42, 6]. In this case, we wish to construct an orthogonal basis for a space of scalar functions over a manifold. In practice, this reduces to computing the eigenfunctions of the Laplace-Beltrami Operator. In Section 4, we will see that this function basis is particularly well-adapted to

the geometry of the underlying manifold on which the function space is defined. First, we develop preliminary tools in the study of manifolds and functional analysis in Section 2. In Section 3 we discuss and implement discrete methods, and in 5 we discuss numerical methods for the Laplace-Beltrami eigenfunction problem.

2 Theoretical Background

In this section, we cover preliminary technical details. These include basic definitions and objects in the study of manifolds, function spaces and partial differential equations. Definitions and notation regarding manifolds are drawn primarily from [43] and [52], while those pertaining to function spaces and differential equations are drawn from [37]. Following this we will review a few foundational manifold learning methods in order to motivate our analysis of Laplace-Beltrami.

2.1 Manifolds and Local Coordinates

The definitions and notation of this section are drawn primarily from [43] and [52]. We will denote $\mathbb{M}$ as a *topological manifold*. That is, $\mathbb{M}$ is a Hausdorff, second-countable, and locally Euclidean topological space. For this work, we further assume that $\mathbb{M}$ is *compact*, *smooth*, equipped with a Riemannian metric and isometrically embedded in a Euclidean space $\mathbb{R}^n$. In fact, a famous result of Nash [35] shows that, given any Riemannian manifold, such a distance-preserving embedding always exists.

To say that $\mathbb{M}$ is *locally Euclidean* of dimension n is to say that for any point x of $\mathbb{M}$ there exists a neighborhood $\mathcal{U}$ of x and a homeomorphism $\phi : \mathcal{U} \to \mathcal{V}$ for which $\mathcal{V}$ is an open subset of $\mathbb{R}^n$. It is common to refer to the pair $(\mathcal{U}, \phi)$ as a *chart*, the set $\mathcal{U}$ a *coordinate neighborhood*, and the map ϕ as a *coordinate map*. By compact, we mean that any collection of coordinate neighborhoods which cover $\mathbb{M}$ will always admit a finite sub-collection which also covers it. On the other hand, the second-countable condition means that a *basis* for the topology on $\mathbb{M}$ is a countable collection. The Hausdorff condition ensures that sequence limits in $\mathbb{M}$ are unique. From a topological perspective, an embedding of $\mathbb{M}$ in a Euclidean space is a continuous and injective mapping $\varphi : \mathbb{M} \to \mathbb{R}^n$ whose restriction $\tilde{\varphi} : \mathbb{M} \to \varphi(\mathbb{M})$ is a homeomorphism.

For any coordinate neighborhood $\mathcal{U}$ of a point x in $\mathbb{M}$, the coordinate map $\phi : \mathcal{U} \to \mathbb{R}^n$ enables one to perform computations in a local neighborhood of the manifold $\mathbb{M}$ as if they were being performed in the Euclidean space $\mathbb{R}^n$. It is common to denote the standard coordinates in $\mathbb{R}^n$ as r^i for $1 \leq i \leq n$. Given a chart $(\mathcal{U}, \phi : \mathcal{U} \to \mathbb{R}^n)$, we denote the *local coordinates* on $\mathcal{U}$ to be the component functions $x^i \equiv r^i \circ \phi : \mathcal{U} \to \mathbb{R}$. Therefore, given any point p of $\mathbb{M}$, we express this point in a Euclidean space as the vector $\phi(p) = (x^1(p), \ldots, x^n(p)) \in \mathbb{R}^n$. Alternatively, one can view r^i as being analogous to the coordinate projection $\pi_i : \mathbb{R}^n \to \mathbb{R}$ defined by $\pi_i(\mathbf{x}) = x_i$.

2.2 Differentiation of Scalar Functions on Manifolds

We now have the machinery in place to begin formulating familiar notions of differential calculus for scalar functions on manifolds. Any two charts $(\mathcal{U}, \phi : \mathcal{U} \to \mathbb{R}^n)$ and $(\mathcal{V}, \psi : \mathcal{V} \to \mathbb{R})$ of $\mathbb{M}$ are called C^∞ *compatible* when the *transition maps*

$$\phi \circ \psi^{-1} : \psi(\mathcal{U} \cap \mathcal{V}) \to \phi(\mathcal{U} \cap \mathcal{V}) \quad \text{and} \quad \psi \circ \phi^{-1} : \phi(\mathcal{U} \cap \mathcal{V}) \to \psi(\mathcal{U} \cap \mathcal{V})$$

are C^∞, or equivalently that each of the above vector-valued functions on $\mathbb{R}^n$ have continuous derivatives of all orders. With this, we call a collection of charts $\{(\mathcal{U}_\alpha, \phi_\alpha)\}$ on $\mathbb{M}$ an *atlas* when the coordinate maps ϕ_α of the collection are pairwise C^∞ compatible and the coordinate neighborhoods $\mathcal{U}_\alpha$ cover $\mathbb{M}$. Such an atlas is *maximal* when it is not contained in a larger atlas, and $\mathbb{M}$ together with a maximal atlas is called *smooth*. Note that $\mathbb{M}$ may be referred to as a *differentiable* manifold when the transition maps are k times continuously differentiable, although we will take $\mathbb{M}$ to be smooth.

Using the notion of coordinate charts, we can define a C^∞ smooth function on a smooth manifold $\mathbb{M}$, and formalize what it means to differentiate such a function. For such $\mathbb{M}$ with dimension n say that $f : \mathbb{M} \to \mathbb{R}$ is C^∞ at a point p of $\mathbb{M}$ if there exists a chart $(\mathcal{U}, \phi)$ with $p \in \mathcal{U}$ such that $f \circ \phi^{-1} : \phi(\mathcal{U}) \to \mathbb{R}$ is C^∞ at $\phi(p)$. The function f is a C^∞ function when this condition holds for each $p \in \mathbb{M}$. In fact, smoothness at p is independent of a choice of chart $(\mathcal{U}, \phi)$ since for any other chart $(\mathcal{V}, \psi)$ we have

$$f \circ \psi^{-1} = (f \circ \phi^{-1}) \circ (\phi \circ \psi^{-1}),$$

and thus the C^∞ smoothness of $f \circ \phi^{-1}$ implies that the same property holds for each $\psi(p) \in \psi(\mathcal{U} \cap \mathcal{V})$. For such a function $f : \mathbb{M} \to \mathbb{R}$ over smooth $\mathbb{M}$ we define its i^{th} *partial derivative* as the function $\frac{\partial f}{\partial x^i} : \phi(\mathcal{U}) \to \mathbb{R}$ such that for $p \in \mathbb{M}$ we have

$$\frac{\partial f}{\partial x^i}(p) = \frac{\partial (f \circ \phi^{-1})}{\partial r^i}(\phi(p)).$$

2.3 Riemannian Manifolds

In Section 2.1, the term *Riemannian manifold* was not explicitly defined, which we address in the present section. This is a key detail which allows the discrete methods of Section 3 to be generalized to the manifold setting. At a high level, a Riemannian manifold is a smooth topological manifold equipped with a *Riemannian metric*[1] inducing a norm, thus allowing one to make sense of the local geometry of the manifold. The definitions and notation here are drawn from [43] and [52].

For any point p of a manifold $\mathbb{M}$ we define the *tangent space* as the vector space whose span contains all vectors tangent to $\mathbb{M}$ at p, which we denote as $\mathcal{T}_p\mathbb{M}$. On this vector space, a Riemannian metric assigns a positive definite bilinear form $g_p : \mathcal{T}_p\mathbb{M} \times \mathcal{T}_p\mathbb{M} \to \mathbb{R}$ to p. This induces a norm $\|\cdot\|_p : \mathcal{T}_p\mathbb{M} \to \mathbb{R}$ defined as $\|v\|_p = \sqrt{g_p(v,v)}$. Suppose that $(x^1, \dots, x^n) : \mathcal{U} \to \mathbb{R}^n$ are the local coordinates in a chart $(\mathcal{U}, \phi)$ of $\mathbb{M}$ where $p \in \mathcal{U}$. Then the set

$$\left\{ \frac{\partial}{\partial x^1}\bigg|_p, \dots, \frac{\partial}{\partial x^1}\bigg|_p \right\}$$

of partial derivatives at p is a basis for $\mathcal{T}_p\mathbb{M}$. It is often useful to consider the family of tangent spaces $\mathcal{T}_x\mathbb{M}$ parameterized by some or all elements x of the topological space $\mathbb{M}$. Such an object is called a *tangent bundle*.[2] Explicitly, the tangent bundle of a differentiable manifold is the disjoint union

$$\mathcal{T}\mathbb{M} = \bigsqcup_{x \in \mathbb{M}} \mathcal{T}_x\mathbb{M} = \bigcup_{x \in \mathbb{M}} \{x\} \times \mathcal{T}_x\mathbb{M} = \{(x, y) : x \in \mathbb{M} \text{ and } y \in \mathcal{T}_x\mathbb{M}\}.$$

Using the basis for $\mathcal{T}_p\mathbb{M}$, we can compute with a Riemannian metric in a local chart. For any $v, w \in \mathcal{T}_p\mathbb{M}$ there exist smooth $\alpha_i, \beta_j : \mathcal{U} \to \mathbb{R}$ such that

$$g_p(v, w) = g_p\left(\sum_{i=1}^n \alpha_i \frac{\partial}{\partial x^i}, \sum_{j=1}^n \alpha_j \frac{\partial}{\partial x^j}\right) = \sum_{i,j} \alpha_i \beta_j g_p\left(\frac{\partial}{\partial x^i}, \frac{\partial}{\partial x^j}\right),$$

which follows from the assumption that g_p is a bilinear form. Therefore, the component functions $g_{ij} : \mathcal{U} \to \mathbb{R}$ of a Riemannian metric take the form

$$g_{ij}(p) = g_p\left(\frac{\partial}{\partial x^i}\bigg|_p, \frac{\partial}{\partial x^j}\bigg|_p\right).$$

This implies that a Riemannian metric is a matrix valued function $G : \mathcal{U} \to \mathbb{R}^{n \times n}$ such that $[G(p)]_{ij} = g_{ij}(p)$. By definition, $G(p)$ is a symmetric positive definite matrix. Many texts refer to defining the family of inner products g_p in a *smooth* way, or that each component function $g_{ij} : \mathcal{U} \to \mathbb{R}$ is smooth. We will suppress the point p from our notation when it can be done without confusion.

Example 1. *For an open set $\mathcal{U} \subset \mathbb{R}^n$ containing a point $\mathbf{p}$ let $(x^1, \dots, x^n)$ denote the standard coordinates on $\mathbb{R}^n$. For any $\mathbf{a}, \mathbf{b} \in \mathcal{T}_p\mathbb{R}^n$ the canonical Euclidean metric is formulated as a Riemannian metric by defining*

$$g_{\mathbf{p}}(\mathbf{a}, \mathbf{b}) = g_{\mathbf{p}}\left(\sum_{i=1}^n a_i \frac{\partial}{\partial x^i}, \sum_{j=1}^n b_j \frac{\partial}{\partial x^j}\right) = \sum_{k=1}^n a_k b_k = \mathbf{a}^\top \mathbf{b}.$$

Appealing to the bilinearity of g_p, we see that

$$\sum_{i,j} a_i b_j g_{ij} = \sum_{k=1}^n a_k b_k,$$

implying that for any $\mathbf{p}$ the local coordinate functions of this Riemannian metric are given by the Kronecker delta $\delta_{ij} = g_{ij}(\mathbf{p})$. Therefore, $G(\mathbf{p}) = \mathbb{I}^{n \times n}$, or the identity matrix.

[1] More generally, a Riemannian metric is a positive definite *metric tensor*. At a high level, a metric tensor endows a purely topological manifold with familiar geometric notions such as angles and distance.

[2] Vector bundles are a more general object of study in algebraic topology and geometry. In the tangent bundle case, we associate to each $x \in \mathbb{M}$ a vector space whose span only contains elements which are tangent to x.

2.4 Function and Inner Product Spaces

For a Riemannian manifold $\mathbb{M}$ with metric G let $C^\infty(\mathbb{M})$ denote the set of all C^∞ smooth scalar functions of the form $f : \mathbb{M} \to \mathbb{R}$ which have *compact support*.[3] We wish to equip this set of functions with an L^2 inner product in a familiar manner. For $\alpha, \beta, \in C^\infty(\mathbb{M})$ and a neighborhood $\mathcal{U} \subset \mathbb{M}$ with coordinates $(x^1, \ldots, x^n) : \mathcal{U} \to \mathbb{R}^n$ we will define

$$\langle \alpha, \beta \rangle = \int_{\mathcal{U} \subset \mathbb{M}} g(\alpha, \beta)\, \sqrt{\det G}\, dx^1 \wedge \cdots \wedge dx^n$$

where $\text{vol}_G = \sqrt{\det G}\, dx^1 \wedge \cdots \wedge dx^n$ is a *volume form* induced by the Riemannian metric G[4] Indeed, we cannot integrate over manifolds in a coordinate independent manner without the notion of a volume form. For a more thorough treatment, see Section 1.2.2 of [43].

Example 2. *The goal of this example is to briefly introduce concepts for which a thorough examination is beyond the scope of this work, and to motivate the fact that our inner product above is well-defined. The volume form* $vol_G = \sqrt{\det G}\, dx^1 \wedge \cdots \wedge dx^k$ *is a certain n-form constructed from the basic* 1*-forms* dx^k *for* $k = 1, \ldots, n$*. Note that* dx^k *are members of the dual vector space* $\mathcal{T}^*\mathbb{M}$*, which we write as* $\mathbb{V}$*. In a chart* $(\mathcal{U}, \phi)$ *with coordinates* $(x^1, \ldots, x^n)$ *on a smooth manifold* $\mathbb{M}$*, a* 1*-form* $\omega : \mathcal{U} \to \mathbb{R}$ *is a linear combination of the kind*

$$\omega = \sum_{k=1}^{n} \alpha_k(x) dx^k$$

where $\alpha_k : \mathcal{U} \to \mathbb{R}$ *are smooth and* $dx^i(\frac{\partial}{\partial x^j}) = \delta_{ij}$*. This gives the intuitive notion that we can integrate with respect to an* n*-form. The operation* $dx^i \wedge dx^j$ *is the multiplication in the exterior algebra* $\bigwedge \mathbb{V}$*. In brief,* $\bigwedge \mathbb{V}$ *is a quotient algebra of the graded tensor algebra* $T(\mathbb{V})$*, which we write*

$$\bigwedge \mathbb{V} = \left(\bigoplus_{k=1}^{n} \mathcal{T}^k(\mathbb{V}) \right) / I$$

where I *is an ideal of* $T(\mathbb{V})$ *and* $T^k(\mathbb{V}) = \bigotimes_{i=1}^{k} \mathbb{V}$*. In fact, one can show that the multiplication* $\wedge$ *is induced by the canonical isomorphism* $T^k(\mathbb{V}) \bigotimes T^l(\mathbb{V}) = T^{k+l}(\mathbb{V})$*, which is true in general for tensor algebras. This means that* vol_G *is in fact an* n*-form. Furthermore, if* $\mathbb{M} = \mathbb{R}^n$ *with standard coordinates we can take the coordinate maps* $\phi : \mathcal{U} \to \mathbb{R}$ *to be the identify* $Id_{\mathcal{U}} : \mathcal{U} \to \mathcal{U}$ *and use our result in Example 1 to write*

$$\begin{aligned}
\langle \alpha, \beta \rangle_{L^2(\mathbb{M}, G)} = \int_{\mathbb{M}} g(\alpha, \beta)\, dvol_G &:= \int_{\mathcal{U} \subset \mathbb{R}^n} g(\alpha, \beta)\, \sqrt{\det \mathbb{I}^{n \times n}}\, dx^1 \wedge \cdots \wedge dx^n \\
&= \int_{\mathcal{U} \subset \mathbb{R}^n} \sum_{k=1}^{n} \alpha_k(x) \beta_k(x)\ dx^1 \wedge \cdots \wedge dx^n \\
&= \int_{\phi(\mathcal{U}) \subset \mathbb{R}^n} \sum_{k=1}^{n} \alpha_k(\mathbf{x}) \beta_k(\mathbf{x})\ d\mathbf{x} \\
&= \int_{\mathcal{U} \subset \mathbb{R}^n} \langle \alpha(\mathbf{x}), \beta(\mathbf{x}) \rangle\ d\mathbf{x},
\end{aligned}$$

which is the standard L^2 *inner product for vector-valued functions over* $\mathbb{R}^n$*. Since we are only considering scalar functions on* $\mathbb{M}$*, then our inner product has the following familiar form*

$$\langle \alpha, \beta \rangle = \int_{\mathbb{M}} \alpha\beta\ dvol_G.$$

A set of functions together with a norm is a *function space*. When this norm is induced from an inner product, the function space is also an *inner product space*. Modulo technical details, the set $C^\infty(\mathbb{M})$ together with the norm induced by the above inner product gives a *Hilbert Space*, or a *complete*[5] inner product space. We denote this Hilbert space as $L^2(\mathbb{M}, G)$ for G being a Riemannian metric. Hilbert spaces have a number of desirable properties that will be useful. For instance, each Hilbert space admits an orthonormal basis $\{\varphi_k\}$ for which each function f of $L^2(\mathbb{M}, G)$ can be written as a linear combination $f = \sum_k c_k \varphi_k$ where the coefficients $\{c_k\}$ can be recovered via an orthogonal projection of the form $f = \sum_k \langle f, \varphi_k \rangle \varphi_k$.

[3] By compact support, we mean a function $f : \mathbb{M} \to \mathbb{R}$ for which the closure of the set $\{x \in \mathbb{M} : f(x) \neq 0\}$ is compact. This makes the induced norm well-defined.

[4] We also assume $\mathbb{M}$ to be connected and oriented, as this avoids technical details that must be handled with care; see [43].

[5] By complete, we mean that all *Cauchy sequences* converge in this inner product space.

2.5 Operators and Laplace-Beltrami in Local Coordinates

An *operator* is a function on a function space. The present work is concerned with the differential operator $\Delta_{\mathbb{M}} f : L^2(\mathbb{M}, G) \to L^2(\mathbb{M}, G)$ given by $\Delta_{\mathbb{M}} f = -\text{div}(\nabla_{\mathbb{M}} f)$, or the *Laplace-Beltrami Operator*. It is tempting to define the gradient of a function $f : \mathbb{M} \to \mathbb{R}$ to be the vector field whose component functions are of the form $\frac{\partial f}{\partial x^i} : \phi(\mathcal{U}) \to \mathbb{R}$. However, this definition implicitly assumes we are working in the standard coordinates on $\mathbb{R}^n$. This leads us to the unfortunate conclusion that defining

$$-\text{div}(\nabla_{\mathbb{M}} f) = \sum_{k=1}^{n} \frac{\partial^2}{\partial (x^k)^2}$$

is not quite correct—our definition must be given in a coordinate free manner. We introduce the shorthand notation $\partial_i = \frac{\partial}{\partial x^i}$, which will be useful here and in ensuing sections.

Definition 1. *For a smooth Riemannian manifold $\mathbb{M}$ with metric G suppose that $f : \mathbb{M} \to \mathbb{R}$ is a smooth function. Define the gradient $\nabla : L^2(\mathbb{M}, G) \to \mathcal{T}\mathbb{M}$ in local coordinates as*

$$\nabla_{\mathbb{M}} f = \sum_{i,j} g^{ij} \partial_i f \partial_j$$

where we take $g^{ij} = g_{ij}(dx^i, dx^j) = [G^{-1}]_{ij}$ without proof.

Proposition 1. *If the local coordinates on $\mathbb{M} = \mathbb{R}^n$ are the standard Euclidean coordinates, then the gradient of any smooth function $f : \mathbb{M} \to \mathbb{R}$ is the standard gradient from multi-variable calculus.*

Proof. This proof is simple. In fact, the most difficult step is proving the identity $g^{ij} = [G^{-1}]_{ij}$, which we take as a given here. Intuitively, we see that the inverse terms g^{ij} of the metric tensor G act to correct differences in our computation that arise from the use of different metrics. Using our result from Example 1 and the fact $g^{ij} = [G^{-1}]_{ij}$, it follows that

$$\nabla_{\mathbb{M}} f = \sum_{i,j} g^{ij} \partial_i f \partial_j = \sum_{i,j} \delta_{ji} \frac{\partial f}{\partial x_i} \partial_j = \sum_{i} \frac{\partial f}{\partial x_i} \partial_i.$$

Note that ∂_i is a basis vector in the tangent bundle $\mathcal{T}\mathbb{M}$, so this vector field is exactly the gradient of f. □

Now we define the divergence in a coordinate-free manner. We noted above that the gradient operator above is a vector field on a manifold, or a smooth function of the form $\mathcal{F} : \mathbb{M} \to \mathcal{T}\mathbb{M}$ which for each $x \in \mathbb{M}$ satisfies $\mathcal{F}(x) = (x, V_x)$ for $V_x \in \mathcal{T}_x M$[6]. We want to define the divergence of $\nabla_{\mathbb{M}} f$ so that it inherits the same properties as the standard divergence in multi-variable calculus. In doing this, it would follow that foundational theorems such as Green's Theorem and the Divergence Theorem would generalize to the manifold setting. To this end, we introduce the following proposition.

Proposition 2. *Let $\mathbb{M}$ be a smooth, Riemannian manifold with $\mathcal{F} : \mathbb{M} \to \mathcal{T}\mathbb{M}$ a vector field such that $\mathcal{F} = \sum_i F_i \partial_i$. Suppose that there exists an operator div$\mathcal{F}$ satisfying*

$$\langle -div\mathcal{F}, f \rangle_{L^2(\mathbb{M}, G)} = \langle \mathcal{F}, \nabla_{\mathbb{M}} f \rangle$$

where the inner product on the right hand side is on the tangent bundle $\mathcal{T}\mathbb{M}$. Then

$$div\mathcal{F} = \frac{1}{\sqrt{\det G}} \partial_j \left(\sqrt{\det G} F^j \right).$$

Proof. Note that for this proof we use the Einstein summation notation

$$\sum_i F_i \partial_i = F^i \partial_i.$$

In words, an index that appears as a superscript and a subscript in a given expression will indicate a summation of that expression over the index. Let $\mathcal{U} \subset \mathbb{M}$ be a coordinate neighborhood with local coordinates $(x^1, \ldots, x^n) : \mathcal{U} \to \mathbb{R}^n$. Let $f \in C^\infty(\mathcal{U})$ and $\mathcal{F} = F^i \partial_i$. Note that for each $x \in \mathcal{U}$ the Riemannian metric G induces an inner product on $\mathcal{T}_x^* \mathbb{M}$ of the form

$$\langle \alpha, \beta \rangle = \int_{\mathcal{U}} g_x(\alpha, \beta) \, d\text{vol}_G,$$

which we may extend to each tensor product $\mathcal{T}_x^* \mathbb{M} \otimes \cdots \otimes \mathcal{T}_x^* \mathbb{M}$. Following the intuition from Example 2, we obtain a global inner product on $\mathcal{T}^* \mathbb{M}$ which we write

$$\langle \alpha, \beta \rangle = \int_{\mathbb{M}} \langle \alpha, \beta \rangle \, d\text{vol}_G.$$

[6] In the literature, it is typical to use language from the study of *fiber bundles* and call $\mathcal{F}$ a *section* of the tangent bundle $\mathcal{T}\mathbb{M}$.

In fact, we may define such an inner product identically on $\mathcal{T}\mathbb{M}$ since it is the dual of $\mathcal{T}^*\mathbb{M}$, and thus isomorphic to it. Now, assuming that div$\mathcal{F}$ does in fact satisfy the given form, we see that:

$$\begin{aligned}
\langle \mathcal{F}, \nabla_{\mathbb{M}} f \rangle &= \int_{\mathbb{M}} \langle \mathcal{F}, \nabla_{\mathbb{M}} f \rangle \, d\text{vol}_G \\
&= \int_{\mathcal{U}} \langle F^i \partial_i, g^{kj} (\partial_k f) \partial_j \rangle \, d\text{vol}_G \\
&= \int_{\mathcal{U}} F^i (\partial_k f) g^{kj} \langle \partial_i, \partial_j \rangle \, d\text{vol}_G \\
&= \int_{\mathcal{U}} F^i (\partial_k f) g^{kj} g_{ij} \, d\text{vol}_G \\
&= -\int_{\mathcal{U}} \frac{1}{\sqrt{\det G}} f \cdot \partial_i (\sqrt{\det G} F^i) \, d\text{vol}_G \\
&= \left\langle -\frac{1}{\sqrt{\det G}} \partial_i \left(\sqrt{\det G} F^i \right), f \right\rangle
\end{aligned}$$

This concludes the proof. □

Chaining together the preceding results, we may define the Laplace-Beltrami operator of a function $f : \mathbb{M} \to \mathbb{R}$ in local coordinates as

$$\Delta_{\mathbb{M}} f = -\text{div}(\nabla_{\mathbb{M}} f) = -\frac{1}{\sqrt{\det G}} \sum_j \partial_j \left(\sqrt{\det G} \sum_i g^{ij} \partial_i f \right).$$

If $\mathbb{M} = \mathbb{R}^n$ with standard coordinates, then it follows that $\det G = 1$ and $g^{ij} = \delta_{ji}$. Therefore, this case yields

$$-\Delta_{\mathbb{M}} f = \sum_{i,j} g^{ij} \frac{\partial}{\partial x^i} \frac{\partial}{\partial x^j} f = \sum_i \frac{\partial^2 f}{\partial x_i^2},$$

which is exactly the standard Laplacian from multi-variable calculus. We will see in Section 5 that representing the operators $\nabla_{\mathbb{M}}$ and $\Delta_{\mathbb{M}}$ in local coordinates will be useful for constructing numerical algorithms.

To conclude this section, we analyze the condition $\langle -\text{div}\mathcal{F}, f \rangle_{L^2(\mathbb{M},G)} = \langle \mathcal{F}, \nabla_{\mathbb{M}} f \rangle$. This means that $-$div is the *adjoint* of $\nabla_{\mathbb{M}}$. More generally, if $L : X_1 \to X_2$ is an operator on inner product spaces X_1, X_2 then the adjoint of the operator L is defined to be the unique *linear*[7] operator $L^* : X_2 \to X_1$ such that $\langle\langle Lx_1, x_2 \rangle\rangle = \langle x_1, L^* x_2 \rangle$ where the inner products $\langle\langle \cdot \rangle\rangle$ and $\langle \cdot \rangle$ are not necessarily the same. The operator $L : X \to X$ is called *self-adjoint* when $\langle\langle Lx_1, x_2 \rangle\rangle = \langle x_1, Lx_2 \rangle$. There are two facts about self-adjoint operators which will be useful for our ensuing discussion, and their proofs are short enough to include here.

Fact 1. *Self-adjoint operators have orthogonal eigenfunctions.*

Proof. Suppose that $f, g \in X$ are eigenfunctions of the operator $L : X \to X$ such that $L^* = L$. It follows that we can write $Lf = \lambda f$ and $Lg = \mu g$ for $\lambda \neq \mu \in \mathbb{R}$. It follows that

$$\langle Lf, g \rangle = \langle f, Lg \rangle \iff \lambda \langle f, g \rangle = \mu \langle f, g \rangle \iff (\lambda - \mu) \langle f, g \rangle = 0.$$

This can only hold when $\langle f, g \rangle = 0$. □

Fact 2. *Self-adjoint operators have real eigenvalues.*

Proof. To do this, we need to consider a *complex* inner product space X for which $L : X \to X$ satisfies $L^* = L$. Suppose that $f \in X$ is an eigenfunction of L so that $Lf = \lambda f$ for some $\lambda \in \mathbb{C}$. Via *sesquilinearity* we have

$$\lambda \|f\|^2 = \langle \lambda f, f \rangle = \langle Lf, f \rangle = \langle f, Lf \rangle = \langle f, \lambda f \rangle = \overline{\lambda} \|f\|^2.$$

It therefore follows that $\overline{\lambda} = \lambda$. Thus $\lambda \in \mathbb{R}$. □

3 Related Work on Discrete Manifold Methods

Now that we've established the theoretical foundation of calculus on manifolds, we're ready to consider discrete cases. Just as the Laplacian on $L^2(\mathbb{R}^n)$ motivated the previous section, the following section will be motivated by the *Graph Laplacian*. For a graph with adjacency matrix A and degree matrix D, its Laplacian is the matrix $L = D - A$. In this section, we will review two discrete manifold learning methods which use this object.

[7] The notion of linearity for an operator is defined in the familiar manner.

3.1 Laplacian Eigenmaps

First we analyze the method of Laplacian Eigenmaps [4]. To motivate the method, consider a set of images of the same object, each consisting of $n \times n$ pixels. We can view each image as a data point in $\mathbb{R}^{n^2}$. However, we can assume that the intrinsic dimension of the space of all these images is the number of degrees of freedom of the camera. Ignoring zoom and lighting, position gives us 3 dimensions and therefore 3 degrees of freedom, and likewise with orientation, etc. Thus, we assume that the structure of the image data lies on a lower dimensional manifold embedded in $\mathbb{R}^{n^2}$. We want to find a way to describe this lower dimensional manifold, $\mathbb{M}$. So, given points $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^l$ (in our example, $l = n^2$), we want to find $\mathbf{y_1}, \ldots, \mathbf{y_k} \in \mathbb{R}^m$ ($m << l$) such that each $\mathbf{y_i}$ represents $\mathbf{x_i}$. See Figure 1 for an example.

An algorithm to do this is described in [4], which uses the points $\mathbf{x_1}, \ldots, \mathbf{x_k} \in \mathbb{R}^l$ to construct a graph with k nodes, and edges connecting neighboring points. Methods to construct such a graph with suitable weights matrix W are discussed in [4]. Once we establish a connected graph, we solve the generalized eigenvalue problem $Lf = \lambda Df$ for $D_{ii} = \sum_j W_{ji}$ and $L = D - W$ the graph Laplacian with adjacency matrix replaced by the weight matrix, W. Supposing $0 = \lambda_0 \leq \lambda_1 \leq \cdots \leq \lambda_{k-1}$ are the generalized eigenvalues with corresponding eigenfunctions $f_0, \ldots, f_{k-1}$, we get an embedding $\mathbf{x_i} \mapsto (f_1(i), \ldots, f_m(i)) = \mathbf{y_i}$ that optimally describes the manifold $\mathbb{M}$ on which our data lies.

A Laplacian Eigenmap Embedding of the Swiss Roll

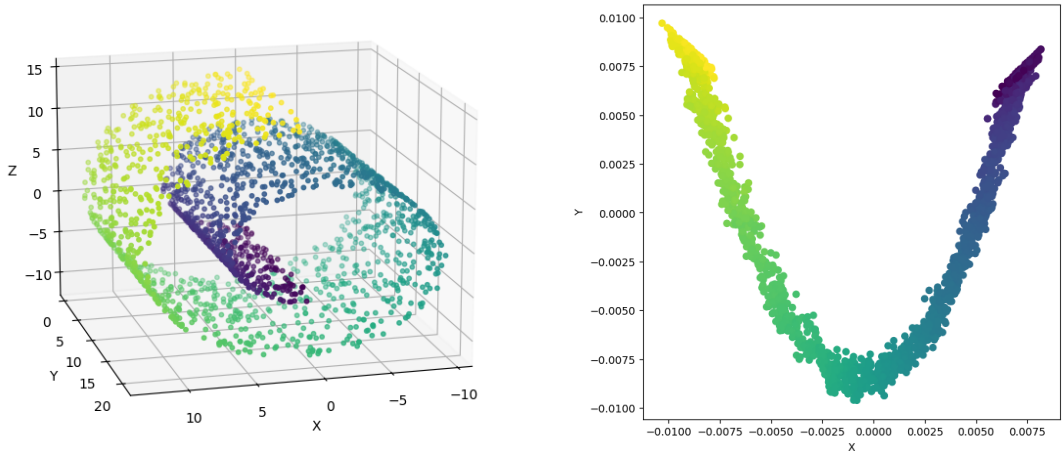


Figure 1: An embedding in $\mathbb{R}^2$ **(right)** of the *Swiss Roll* in $\mathbb{R}^3$ **(left)** with $n = 2000$ and computed via a Laplacian Eigenmap using a Gaussian heat kernel to construct the weighted adjacency matrix. Notice that points which are close on the Swiss Roll map to points which are close on the embedding.

3.2 Local Linear Embedding

Local linear embedding (LLE) [44] is a method to construct a low-dimensional representation of high-dimensional data that preserves local linearity. Since it is not a globally linear method, it is able to work effectively on data that lies on a nonlinear manifold; see Figure 2. The method involves, for each data point, finding its n nearest neighbors and assembling a weight matrix W by choosing W_{ij} to minimize $\sum_{i=1}^{\ell} \|x_i - \sum_{j=1}^{n} W_{ij} x_{i_j}\|$ for all data points $\mathbf{x}_i \in \mathbb{R}^k$. From there, the lower-dimensional data is represented by the eigenvectors corresponding to the k lowest eigenvalues of $E = (I - W)^\top (I - W)$. Under conditions, if E is considered an operator of a function f defined over the data, then $Ef = \frac{1}{2}\Delta^2_{\mathbb{M}} f$. Since the eigenvectors of $\frac{1}{2}\Delta^2_{\mathbb{M}}$ and $\Delta_{\mathbb{M}}$ are the same, the eigenvector problem of E becomes an eigenfunction problem of $\Delta_{\mathbb{M}}$. Observe that we can approximate $[(I - W)f]_i$ as

$$[(I - W)f]_i \approx -\frac{1}{2} \sum_j W_{ij} (x_i - x_{i_j})^\top H (x_i - x_{i_j}),$$

and if a convex combination of $\sqrt{\alpha_i} v_i$s form an orthonormal basis, then $\sum_j W_{ij} v_j^\top H v_j = \Delta_{\mathbb{M}} f$ [4].

Interestingly, while both Laplacian Eigenmaps and LLE can be traced to eigenfunction problems of the Laplace-Beltrami operator, they arrive at very different solutions in practice (See Fig. 1 and Fig. 2). Both methods preserve locality, which we will strengthen mathematically in 4.1, but "unroll" the data very differently.

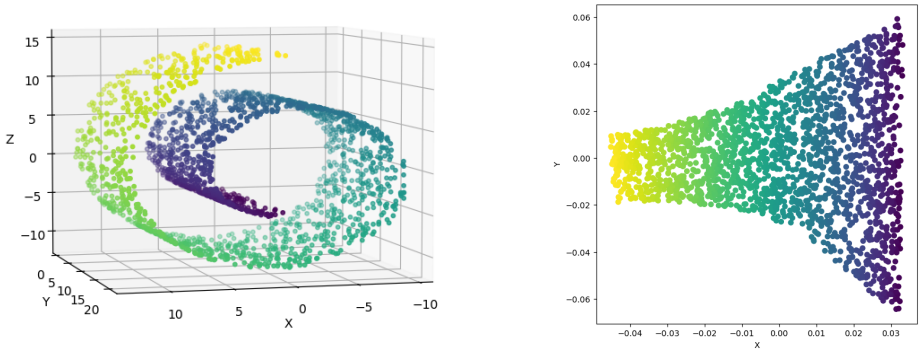


Figure 2: A Locally Linear Embedding in $\mathbb{R}^2$ (right) of the *Swiss Roll* in $\mathbb{R}^3$ (left). Notice that the local manifold structure of the surface is preserved when embedded in the plane.

3.3 Persistent Homology

Around the same time that works such as [44] and [50] introduced nonlinear dimensionality reduction, the problem of estimating geometric and topological invariants of a manifold from a point cloud sample was studied extensively in the literature. From these methods, *persistent homology* has emerged as an increasingly popular choice for practitioners [62, 19], and variants of this computation have been used to develop manifold learning methods [15]. To compute persistent homology of a point cloud, one builds a hierarchical *complex* from the point cloud induced by an increasing parameter. In the simplest case, this parameter is simply a pairwise distance threshold. In practice, this is most often a filtered *simplicial complex*. In fact, such an object and its hierarchical structure can be represented by constructing so-called *boundary matrix* whose row and column indices have an order respecting the hierarchical structure of the complex. This matrix is factored [11], and from this we construct a *graded persistence module* [62]. A number of results have been provided regarding the stability of this computational paradigm [10, 31, 8].

A result of Gabriel [22] implies that each persistent homology module over a finite complex may be written as a direct sum of the form

$$H_*(F_\bullet X) \cong \bigoplus_{\xi} \mathbb{I}\langle b_\xi, d_\xi \rangle.$$

In words this means that, across all dimensions, the homology of the filtered complex X can be expressed as the direct sum of *birth* and *death* interval modules corresponding to each homological feature, ξ. In the literature, it is common to refer to these homological features as *cycles*. By birth and death, we mean the parameter value at which the feature emerges and vanishes in the filtration. This gives the intuitive notion that the true homological signal of the data is captured by features for which $d_\xi - b_\xi$ is large. Like the Laplacian methods that we study in this work, persistent homology understands local features of a manifold in order to build a global geometric picture; see Figure 3.

4 Properties of Laplace-Beltrami

Following from the discrete cases, we're ready to generalize to the continuous case using the Laplace-Beltrami Operator. In this section, we are particularly interested in understanding the properties of the Laplace-Beltrami Operator which make its eigenfunctions well-adapted to the geometry of the underlying manifold.

4.1 Preservation of Locality

To start, we consider the problem of looking for functions $f : \mathbb{M} \to \mathbb{R}$ such that points close together on $\mathbb{M}$ are mapped to points close together in $\mathbb{R}$. As in Section 2.1, we assume $\mathbb{M}$ is a smooth, compact, m-dimensional Riemannian manifold embedded in $\mathbb{R}^l$.

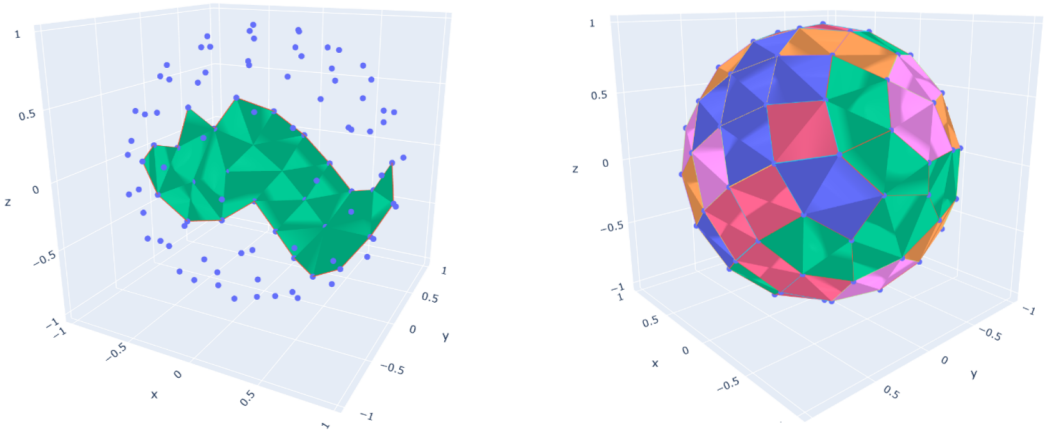
Consider neighboring points $\mathbf{x}$,$\mathbf{z} \in \mathbb{M}$. They map to $f(\mathbf{x})$, $f(\mathbf{z})$. We claim that

$$|f(\mathbf{x}) - f(\mathbf{z})| \leq d_{\mathbb{M}}(\mathbf{x}, \mathbf{z}) \|\nabla f(\mathbf{x})\| + o(d_{\mathbb{M}}(\mathbf{x}, \mathbf{z})).$$

To show this, we use a proof from [4], which takes advantage of a *geodesic curve* $c(t)$ parameterized by length between $\mathbf{x} = c(0)$ and $\mathbf{z} = c(l)$. Letting $l = d_{\mathbb{M}}(\mathbf{x}, \mathbf{z})$, we know by the Fundamental Theorem of Calculus that

$$f(\mathbf{z}) - f(\mathbf{x}) = f(c(l)) - f(c(0)) = \int_0^l \frac{d}{dt} f(c(t))\, dt$$

Estimating $\mathbb{S}^2$ from a point cloud via Persistent Homology



(a) A single homological cycle **(orange)** together with its bounding chain **(green)**.

(b) The collection of bounding chains associated with all homological cycles.

Figure 3: A point cloud sampled from $\mathbb{S}^2$ with local **(a)** and global **(b)** structure of the manifold determined using persistent homology. The two dimensional features along the surface are often referred to as *bounding chains*, while the one dimensional features enclosing these bounding chains are the homological cycles.

where $c'(t)$ is the tangent vector at $c(t)$. So, $\frac{d}{dt}f(c(t)) = \frac{df}{dc(t)}c'(t)$, which implies that, on $\mathbb{M}$, we have $\frac{d}{dt}f(c(t)) = df_{c(t)}(c'(t))$. Solving out, we get

$$f(\mathbf{z}) = f(\mathbf{x}) + \int_0^l df_{c(t)}(c'(t))\,dt = \int_0^l \langle \nabla f(c(t)), c'(t)\rangle\,dt + f(\mathbf{x})$$

where the integral term on the right hand side is the integral of the directional derivative on $\mathbb{M}$ in the direction $c'(t)$. It therefore follows that

$$|f(\mathbf{z}) - f(\mathbf{x})| = \left| \int_0^l \langle \nabla f(c(t)), c'(t)\rangle\,dt \right|.$$

Applying Cauchy-Schwartz with $\|c'(t)\| = 1$, we get

$$|\langle \nabla f(c(t)), c'(t)\rangle| \leq \|\nabla f(c(t))\| \cdot \|c'(t)\| \implies |\langle \nabla f(c(t)), c'(t)\rangle| \leq \|\nabla f(c(t))\|.$$

Using a Taylor approximation at $t = 0$ with smoothness of f and the fact that $c(0) = \mathbf{x}$ gives us

$$\|\nabla f(c(t))\| = \|\nabla f(c(0))\| + t\frac{d}{dt}\|\nabla f(c(t))\| + O(t^2) = \|\nabla f(\mathbf{x})\| + O(t)$$

$$\implies |f(\mathbf{z}) - f(\mathbf{x})| \leq \int_0^l \|\nabla f(\mathbf{x})\|\,dt = l\|\nabla f(\mathbf{x})\| + O(l)$$

following from the fact that $f(\mathbf{x})$ is not a function of t. Since $\mathbb{M} \subset \mathbb{R}^l$, we know that $d_{\mathbb{M}}(\mathbf{x}, \mathbf{z}) = \|\mathbf{x} - \mathbf{z}\|_{\mathbb{R}^l} + o(\|\mathbf{x} - \mathbf{z}\|_{\mathbb{R}^l})$. Note that $l = d_{\mathbb{M}}(\mathbf{x}, \mathbf{z})$, resulting in the inequality

$$|f(\mathbf{z}) - f(\mathbf{x})| \leq d_{\mathbb{M}}(\mathbf{x}, \mathbf{z})\|\nabla f(\mathbf{x})\| + o(d_{\mathbb{M}}(\mathbf{x}, \mathbf{z})).$$

From this we can conclude that $\|\nabla f\|$ provides us with an upper bound on how far away $\mathbf{x}$ and $\mathbf{z}$ are mapped to under f, so in other words an estimate of preservation of locality. In general, we're looking for an f that best preserves locality on average, leading to the minimization problem

$$\underset{\|f\|_{l^2(\mathbb{M})}=1}{\arg\min} \int_{\mathbb{M}} \|\nabla f(\mathbf{x})\|^2$$

with the unit-norm constraint intended to avoid trivial solutions $f = 0$. [8]

[8] The term $\int_{\mathbb{M}} \|\nabla f(\mathbf{x})\|^2$ is called the "Dirichlet energy" and measures the total variation of a function in a space, so minimizing it enforces smoother eigenfunctions.

We consider the graph version of this problem: data points turn into nodes on the graph with edges weighted by matrix W, which is chosen advisedly as described in [4]. Defining $f_i = f(x_i)$, we get that $\int_{\mathbb{M}} \|\nabla f(\mathbf{x})\|^2$ is analogous to $f^T L f = \frac{1}{2}\sum_{i,j}(f_i - f_j)^2 W_{ij}$ where L is the discrete graph Laplacian described in 3.1. Thus, the above minimization problem with appropriate constraints leads (in the continuous setting) to the eigenvalue problem for the Laplace–Beltrami operator, $\Delta_{\mathbb{M}}$.

As discussed in 2.2, $\Delta_{\mathbb{M}} f = -\text{div}(\nabla_{\mathbb{M}} f)$, and by Stokes' Theorem, $-$div and ∇ are adjoint operators—i.e. for a vector field X we have $\int_{\mathbb{M}} \langle X, \nabla f\rangle = -\int_{\mathbb{M}} \text{div}(X) f$. Therefore, we conclude that

$$\int_{\mathbb{M}} \|\nabla f\|^2 = \int_{\mathbb{M}} f \Delta_{\mathbb{M}} f.$$

More precisely, $\Delta_{\mathbb{M}}$ is positive semi-definite (see 4.2), so a function f that minimizes $\int_{\mathbb{M}} \|\nabla f\|^2$ is an eigenfunction of $\Delta_{\mathbb{M}}$. Denote the eigenvalues of $\Delta_{\mathbb{M}}$ as $0 = \lambda_0 \leq \lambda_1 \leq \dots$ and let f_i be the eigenfunction corresponding to λ_i.

One consideration is that $\lambda_0 = 0$, so $\Delta_M f_0 = 0$, and thus $\nabla f_0(\mathbf{x}) = 0$ for each $\mathbf{x}$. Because of this, we require all other f_i to be orthogonal to f_0, meaning that $\int_{\mathbb{M}} f_i(\mathbf{x}) f_0(\mathbf{x})\, d\mathbf{x} = 0$. We discuss orthogonality and other constraints on the eigenfunctions in 4.2. Therefore, we can choose the first m eigenfunctions $f_1, \dots f_m$ orthogonal to each other and to f_0 and map $\mathbf{x} \mapsto (f_1(\mathbf{x}), \dots, f_m(\mathbf{x})) \in \mathbb{R}^m$. Since each f_i is smooth, this embedding preserves locality in the sense of minimizing Dirichlet energy. In particular, since we are minimizing $\int_{\mathbb{M}} \|\nabla f\|^2$ over all of $\mathbb{M}$, the Laplace-Beltrami Operator promotes smooth eigenfunctions that capture the global geometry of $\mathbb{M}$.

In the discrete case of Laplacian Eigenmaps, it has been shown in [5] that eigenvectors of the Graph Laplacian converge to eigenfunctions of the Laplace-Beltrami Operator on $\mathbb{M}$. As a result of this convergence, the discrete case of Laplacian Eigenmaps allows us to capture a consistent approximation of the global geometry of $\mathbb{M}$ in practical use.

4.2 Self-Adjointness and Positive Semi-Definite

The Laplace-Beltrami Operator is easily shown to be self-adjoint under the definition $\Delta_{\mathbb{M}} = -\text{div}(\nabla f)$. By our discussion in Section 2.5, we may use the fact that the adjoint of the gradient, ∇, is the negative divergence, $-$div under the L^2 inner product. From there, using properties of inner products, we can see that for any functions f, g that

$$\langle \Delta_{\mathbb{M}} f, g\rangle = \langle -\text{div}(\nabla f), g\rangle = \langle \nabla f, \nabla g\rangle = \langle f, -\text{div}(\nabla g)\rangle = \langle f, \Delta_{\mathbb{M}} g\rangle.$$

Therefore, the Laplace-Beltrami Operator is self-adjoint. By Fact 2, this ensures that the eigenvalues of the operator are real. Thus, considering the eigenfunctions associated with the first m eigenvalues (excluding λ_0) makes sense—ordering the eigenvalues is only well-defined if they are real. On the other hand, Fact 1 implies that the eigenfunctions of the Laplace-Beltrami Operator are orthogonal. This fact about the eigenfunctions may be leveraged for computational purposes; see our application with *Galerkin projections* [51] in Section 5.

Furthermore, we can show that the Laplace-Beltrami Operator is positive semi-definite:

$$\langle \Delta_{\mathbb{M}} f, f\rangle = \langle -\text{div}(\nabla f), f\rangle = -\int_{\mathbb{M}} f \cdot \text{div}(\nabla f) = -\int_{\mathbb{M}} \text{div}(f\nabla f) + \int_{\mathbb{M}} |\nabla f|^2 = \int_{\mathbb{M}} |\nabla f|^2 \geq 0$$

This follows from the Divergence Theorem: $\mathbb{M}$ is Riemannian compact so $\int_{\mathbb{M}} \text{div}(f\nabla f) = \int_{\partial \mathbb{M}} (f\nabla f)\cdot \mathbf{n} dS = 0$. As a result of this, the eigenvalues of the Laplace-Beltrami Operator are not only real but also nonnegative, so smaller eigenvalues truly correspond to eigenfunctions with less fluctuation, allowing us to ensure preservation of locality.

4.3 Nodal Domains

For a more thorough overview, we direct the reader to [26, 32]. Given an eigenfunction $v : \mathbb{M} \to \mathbb{R}$ of the Laplace-Beltrami operator $\Delta_{\mathbb{M}}$ with Dirichlet, Neumann or Robin boundary condiitons, its *nodal set* is given by

$$\mathcal{N} = \{x \in \mathbb{M} : v(x) = 0\}.$$

In fact, nodal sets have been studied dating back to the work of Ernst Chladni's 1787 book "Discoveries concerning the Theories of Music". In this work, Chladni accounts observations made while vibrating a thin metal plate with a bow and spreading sand over it. Sand accumulates in stationary zones, and gives rise to surprisingly complex patterns. In fact, Chladni's experiment was modeling stationary waves, which can be understood by the spatial component of the *Helmholtz equation* $\Delta f = -\lambda f$. Thus, this problem is identical to computing the eigenfunctions of the Laplace-Beltrami operator. Nodal sets have been studied by Lord Rayleigh [39], and Courant and Sturm [12]. Furthermore, for any compact Riemannian manifold, $\mathbb{M}$, the eigenfunctions of $\Delta_{\mathbb{M}}$ may be used to construct an orthonormal basis for $L^2(\mathbb{M})$, and this basis of eigenfunctions is well-adapted to the underlying manifold, in that:

1. The nodal set of any eigenfunction partitions the underlying manifold into *nodal domains*, and, by ordering the eigenvalues of $\Delta_{\mathbb{M}}$ such that $\lambda_0 = 0 < \lambda_1 \leq \lambda_2 \leq \cdots$, then the eigenfunction associated with λ_n can have no more than n nodal domains.
2. The nodal sets are curves intersecting at constant angles.

4.4 Connection to the Heat Equation

Another conceptualization of the Laplace-Beltrami Operator comes from the heat equation, a classic problem in partial differential equations [4]. Denote $\Delta_{\mathbb{M}}$ as the Laplace-Beltrami operator, so $\Delta_{\mathbb{M}} f = -\text{div}\nabla(f)$. The heat equation is defined as the problem $\left(\frac{\partial}{\partial t} + \Delta_{\mathbb{M}}\right) u = 0$, where $u(t,x)$ is the heat distribution at time t and $f : \mathbb{M} \to \mathbb{R}$ is the initial heat distribution. The solution to this partial differential equation is $u(t,x) = \int_{\mathbb{M}} H_t(x,y) f(y)$, where H_t is the heat kernel or Green's function for the equation. Since the solution to the PDE is known, we can use it to discern the Laplace-Beltrami Operator on a function, and find that

$$\Delta_{\mathbb{M}}(f(x)) = -\Delta_{\mathbb{M}} u(x,0) = -\left(\frac{\partial}{\partial t}\left[\int_{\mathbb{M}} H_t(x,y) f(y)\right]\right)_{t=0}.$$

This version of the Laplace-Beltrami operator allows us to use well-known approximations from partial differential equations to estimate the Laplace-Beltrami Operator's action on a function. In some coordinate systems, $H_t \approx (4\pi t)^{-m/2} e^{-\frac{\|x-y\|^2}{4t}} (\phi(x,y) + O(t))$, the Gaussian, where $\phi(x,y)$ is smooth with $\phi(x,x) = 1$. Thus, when x and y are close to each other and t is small,

$$H_t \approx (4\pi t)^{-\frac{m}{2}} e^{-\frac{\|x-y\|^2}{4t}}.$$

Furthermore, as $t \to 0$ we have $H_t \to \delta$, so $\lim_{t\to 0} \int_{\mathbb{M}} H_t(x,y) f(y) = f(x)$. From the definition of the derivative,

$$\begin{aligned}\Delta_{\mathbb{M}}(f(x)) &= -\left(\frac{\partial}{\partial t}\left[\int_{\mathbb{M}} H_t(x,y) f(y)\right]\right)_{t=0} = -\lim_{t\to 0} \frac{\int_{\mathbb{M}} H_t(x,y) f(y) - \int_{\mathbb{M}} H_0(x,y) f(y)}{t} \\ &= \frac{1}{t}\left(\int_{\mathbb{M}} H_0(x,y) f(y) - \lim_{t\to 0} \int_{\mathbb{M}} H_t(x,y) f(y)\right) \approx \frac{1}{t}\left(f(x) - (4\pi t)^{-\frac{m}{2}} \int_{\mathbb{M}} e^{-\frac{\|x-y\|^2}{4t}} f(y) dy\right).\end{aligned}$$

After substituting data points on $\mathbb{M}$ in for x, this approximation, driven from PDEs, leads to results in dimensionality reduction, as it determines a possible scheme for assigning weights between data points.

4.5 Laplace-Beltrami Operator on $\mathbb{S}^2$

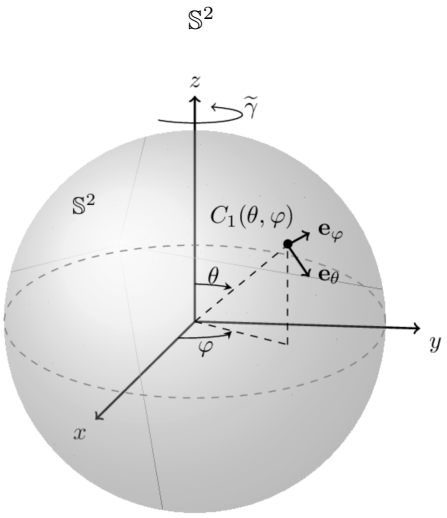


Figure 4: A visualization of $\mathbb{S}^2$ illustrating the colatitude (θ) and the azimuthal angle (φ) found in [23]. These are the general spherical coordinates, with $r = 1$ everywhere since we are working with the unit sphere.

The $\mathbb{S}^2$ manifold, effectively the surface of a sphere in $\mathbb{R}^3$, provides an interesting case of the Laplace-Beltrami operator. For a more thorough review, see [14]. In $\mathbb{R}^3$, points on the manifold are defined as $x = \cos\varphi \sin\theta, y = \sin\varphi \sin\theta$, and $z = \cos\theta$, where θ is the colatitude and φ the azimuthal angle (Fig. 4). These angles are known generally to define spherical coordinates, and since we are working with the unit circle, the radius is always 1. A well-defined function f on $\mathbb{S}^2$ has the properties $f(\theta,\varphi) = f(\theta,\varphi+2\pi)$ (it must follow the sphere) and $f(0,\varphi) = f(\pi,\varphi)$ (the north and south pole must match). If we define the area element on $\mathbb{S}^2$ to be $d\mu$, then $d\mu = \sin\theta d\theta d\varphi$. Thus, if f is well-defined and continuous on $\mathbb{S}^2$, then $\int_{\mathbb{S}^2} f d\mu = \int_0^{2\pi}\int_0^{\pi} f(\theta,\varphi)\sin\theta d\theta d\varphi$. Since $\mathbb{S}^2 \subset \mathbb{R}^3$, we can use

the $\mathbb{R}^n$ definition of the Laplacian to find the Laplace-Beltrami operator on $\mathbb{S}^2$. Recall that $\Delta_{\mathbb{R}^n} = \sum_{i=0}^{n} \frac{\partial^2 f}{\partial x_i^2}$, so $\Delta_{\mathbb{R}^3} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} + \frac{\partial f}{\partial z}$. Then, by the Chain Rule, $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial \theta}\frac{\partial \theta}{\partial x} + \frac{\partial f}{\partial \varphi}\frac{\partial \varphi}{\partial x}$, with the same applying to y and z. After some calculus, we arrive at the definition of the Laplace-Beltrami Operator (with the operator positive in this case)

$$\Delta_{S^2} = \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}.$$

The eigenfunctions of the Laplace-Beltrami operator are the spherical harmonics of f, the spherical analogue of the Fourier series expansion, while each eigenvalue has a basis of eigenfunctions that are finite dimensional [14].

5 Discussion: Numerical Methods

We have done much work to provide a rigorous definition of Laplace-Beltrami starting only with the definition of a purely topological manifold. In doing so, we streamlined the presentation of key results and properties in Section 4. However, we are also motivated by algorithms and applications. In Section 3 we described two discrete Laplacian manifold methods and a related topological method for machine learning and dimensionality reduction. In Section 4.1, the argument detailed from [4] motivates Laplace-Beltrami as a generalization of these methods to a continuous setting. In this section, we exemplify these ideas with applications. First, we review relevant literature.

5.1 Literature

The Laplace-Beltrami operator first garnered attention from researchers in dimensionality reduction and manifold learning as a unifying theoretical foundation for Laplacian manifold methods [6]. This followed from the the study of Laplacian Eigenmaps [4] and their convergence [5]. Recently, the authors of [58] extend this framework to Finsler manifolds [21], where the tangent spaces of the underlying manifold are equipped with a norm that need not be induced by an inner product. Much attention has been given to efficient and accurate computation of the Laplace-Beltrami eigenprobem since the eigenfunctions are well-adapted to the geometry of the underlying manifold. Indeed, the operator plays a foundational role in *geometry processing* tasks such as *mesh parameterization*, *shape analysis*, and the construction of *shape descriptors*, for example [58, 17, 32, 41, 42, 55, 49].

In practice, Laplacian methods reduce to an eigenvalue problem in the discrete cases of Section 3 and an eigenfunction problem in the continuous case. The most classical method in the continuous case is the *cotangent method* of [38], which is a *Finite Element Method* (FEM) over a triangular mesh that uses a discretized Laplace-Beltrami operator derived via Green's Theorem. See Section 5.2 for an example with the heat equation and Section 5.3 for a similar geometry processing example. The cotangent scheme does not provide point-wise convergent estimates in general, as exemplified by [60, 57]. To this end, Belkin et. al. [7] presented the first method providing a discretized approximate to the Laplace operator on a meshed surface with guarantees of point-wise convergence. In [61] the authors use persistent homology—see section 3.3—as regularization for Laplace-Beltrami manifold learning and show that is produces precise discretizations of the operator.

In [59], the authors present a fast discretized method to compute Laplace-Beltrami eigenproblems via a *linear subspace* method, while [36] presents a method to approximate Laplace-Beltrami eigenvalues together with empirical evidence that it saves on memory and run-time. The orthogonality of the eigenfunctions of the operator makes a *Galerkin* FEM [51] approach viable. See [42] for the original work and Section 5.3 for an example. The spectrum of Laplace-Beltrami eigenvalues may also be learned via a *graph neural network*, as is recently shown in [2]. It is important to note the difference between a discretized Laplace-Beltrami operator and a discrete Laplacian, as is noted in [32]. Here, the author uses a discrete Laplacian as defined in [16] to compute an orthogonal basis for a function space over a surface given by the eigenfunctions of the Lapalce-Beltrami operator.

5.2 Heat Equation Application

To demonstrate how the Laplace-Beltrami operator can be used to numerically solve the heat equation on a manifold, consider the Möbius strip of width w, parameterized in $\mathbb{R}^3$ as $\mathbf{r}(u,v) = [(1 + v\cos(\frac{u}{2}))\cos(u), (1 + v\cos(\frac{u}{2}))\sin(u), v\sin(\frac{u}{2})]$ for $u \in [0, 2\pi), v \in [-w, w]$. We first construct a mesh of quads $\begin{bmatrix} i_0 & i_1 \\ i_2 & i_3 \end{bmatrix}$ with $i_0 = in_u + j$, $i_1 = in_u + j + 1$, $i_2 = (i+1)n_u + j$, and $i_3 = (i+1)n_u + j + 1$. These define vertices of the quads in a mesh indexed by $i = 0, \ldots, n_v - 1$, $j = 0, \ldots, n_u$. Because the Möbius strip flips at the seam—the column of quads at $j = n_u - 1$)—the v coordinate flips there: $i_0 = in_u + n_u - 1$, $i_1 = (n_v - 1 - i)n_u$, $i_2 = (i+1)n_u + n_u - 1$, and $i_3 = (n_v - 1 - (i+1))n_u$. Inside of each quad, we can define two triangles (i_0, i_1, i_3) and (i_0, i_3, i_2), giving us a mesh of triangles approximating the Möbius strip $\mathbb{M}$. (See [9] for more detail).

As discussed in 4.4, the heat equation is $u_t = \Delta_{\mathbb{M}} u$. We use the weak form with a test function v in Hilbert space: $\int_{\mathbb{M}} v u_t = \int_{\mathbb{M}} v\Delta_{\mathbb{M}} u$. By Green's Theorem on manifolds, $\int_{\mathbb{M}} v\Delta_{\mathbb{M}} u = -\int_{\mathbb{M}} \nabla_{\mathbb{M}} v \cdot \nabla_{\mathbb{M}} u + \oint_{\partial\mathbb{M}} v(\nabla_{\mathbb{M}} u \cdot \hat{\mathbf{n}})ds$.

The line integral term becomes zero when we impose Neumann boundary conditions (no net flow in or out of $\mathbb{M}$), so $\int_{\mathbb{M}} v\Delta_{\mathbb{M}}u = -\int_{\mathbb{M}} \nabla_{\mathbb{M}}v \cdot \nabla_{\mathbb{M}}u$.

This holds for any test function v in Hilbert space, so we apply a Galerkin projection to a subspace V_h spanned by basis functions $\phi_1, \ldots, \phi_N$ where for $\mathbf{p}_i$ we have $\phi_j(\mathbf{p}_i) = \delta_{ij}$. We're looking for an approximate solution $u_h \in V_h$ with $u_h(\mathbf{x}, t) = \sum_{j=1}^{N} u_j(t)\phi_j(\mathbf{x})$. Namely, u_h should satisfy $\forall v \in V_h$: $\int_{\mathbb{M}} v\frac{\partial u_h}{\partial t} = -\int_{\mathbb{M}} \nabla_{\mathbb{M}}v \cdot \nabla u_h$ [51].

In particular, $V_h = \text{span}(\phi_1, \ldots, \phi_N)$, so set $v = \phi_i$ for some i:

$$\int_{\mathbb{M}} \phi_i \frac{\partial u_h}{\partial t} = \int_{\mathbb{M}} \phi_i \sum_j \dot{u_j}(t)\phi_j = \sum_j \dot{u_j}(t) \int_{\mathbb{M}} \phi_i\phi_j$$

and

$$-\int_{\mathbb{M}} (\nabla_{\mathbb{M}}\phi_i \cdot \sum_j u_j(t)\nabla_{\mathbb{M}}\phi_j) = -\sum_j u_j(t) \int_{\mathbb{M}} (\nabla_{\mathbb{M}}\phi_i \cdot \nabla_{\mathbb{M}}\phi_j)$$

so by the heat equation,

$$\sum_j \dot{u}_j(t) \int_{\mathbb{M}} \phi_i\phi_j = -\sum_j u_j(t) \int_{\mathbb{M}} (\nabla_{\mathbb{M}}\phi_i \cdot \nabla_{\mathbb{M}}\phi_j).$$

Setting $M_{ij} = \int_{\mathbb{M}} \phi_i\phi_j$, $L_{ij} = \int_{\mathbb{M}} (\nabla_{\mathbb{M}}\phi_i \cdot \nabla_{\mathbb{M}}\phi_j)$, we get a first order system of N coupled ordinary differential equations $M\dot{\mathbf{u}} = -L\mathbf{u}$. With a bit more work, we can compute L and M explicitly based on the parameters of our triangle mesh of the manifold with a method described in [38]. With our Möbius strip, we get

$$L_{ij} = \begin{cases} \frac{1}{2}\sum_{T\ni i}(\cot\alpha_i^T + \cot\beta_i^T) & j = i \\ -\frac{1}{2}(\cot\alpha_{ij} + \cot\beta_{ij}) & j \in \mathcal{N}(i) \\ 0 & \text{otherwise} \end{cases} \qquad M_{ij} = \begin{cases} \frac{1}{3}\sum_{T\ni i} A_T & j = i \\ 0 & j \neq i \end{cases} \tag{1}$$

for α_{ij}, β_{ij} the two angles opposite edge (i, j), $\mathcal{N}(i)$ the set of vertices sharing an edge with i, and A_T the area of triangle T.

We can solve this system of ODEs with implicit Euler: $\frac{(\mathbf{u}_{n+1}-\mathbf{u}_n)}{\Delta t} \approx \dot{\mathbf{u}} = -M^{-1}L\mathbf{u}_{n+1}$

$$\mathbf{u}_{n+1} + \Delta t\, M^{-1}L\mathbf{u}_{n+1} = \mathbf{u}_n \tag{2}$$

$$(I + \Delta t\, M^{-1}L)\mathbf{u}_{n+1} = \mathbf{u}_n \tag{3}$$

$$(M + \Delta t\, L)\mathbf{u}_{n+1} = M\mathbf{u}_n \tag{4}$$

and applying LU decomposition to $M + \Delta tL = PL'U$, we get $\mathbf{u}_{n+1} = U^{-1}L'^{-1}P^T M\mathbf{u}_n$. This is our iteration we can apply to approximate $\mathbf{u}$ as time increases. In Figure 5 we see the results in log time (since heat transfer happens very quickly) of a heat spike applied to an initial point on the Möbius strip. The code to compute this and create the visual (Figure 5) can be found in [27].

The connection between the Laplace-Beltrami operator and this method of solving the heat equation is intimate: picking our initial condition $\mathbf{u}_0$, we can decompose it as $\mathbf{u}_0 = \sum_k c_k\phi_k$ where ϕ_k are the generalized eigenfunctions of L ($L\phi_k = \lambda_k M\phi_k$) and $c_k = \phi_k^T M\mathbf{u}_0$ are coefficients. Note that $\lambda_0 = 0$, so the $k = 0$ term in the solution satisfies $c_0 e^{-\lambda_0 t}\phi_0 = c_0\phi_0$, which is constant in time and represents the steady-state heat distribution. This is consistent with conservation of total heat imposed by the Neumann boundary conditions, since no energy escapes through $\partial\mathbb{M}$.

Since L is the Galerkin projection of $\Delta_{\mathbb{M}}$, eigenfunctions of L are the best discrete approximation of eigenfunctions of $\Delta_{\mathbb{M}}$. So indeed, the solution $\mathbf{u}(t) = \sum_{k=0}^{N-1} c_k e^{-\lambda_k t}\phi_k$ is fully determined by the eigenfunctions of L, which by the Galerkin construction converge to eigenfunctions of $\Delta_{\mathbb{M}}$ as the mesh is refined (See [5]).

5.3 Helmholtz Equation and Geometry Processing

In geometry processing, many tasks reduce to the computation of a structured function basis over a manifold or surface. In the literature, the eigenfunctions of the Laplace-Beltrami operator have proven to be a popular choice of function basis, as is noted in Section 5.1. In practice, computing these eigenfunctions is done by computing solutions to the *Helmholtz equation*. For a manifold $\mathbb{M}$ and an element f of the Hilbert space $L^2(\mathbb{M})$, the (homogeneous) Helmholtz operator is given by $\Delta_{\mathbb{M}}f + \lambda f = 0$ for $\lambda \in \mathbb{R}$. In this section we detail a numerical solution method of [42] using a Galerkin FEM approach and a discretized Helmholtz operator derived via Green's formula in a manner similar to that of Section 5.2. As a result, the problem is transformed to a generalized eigenproblem. See Figure 6 for an example.

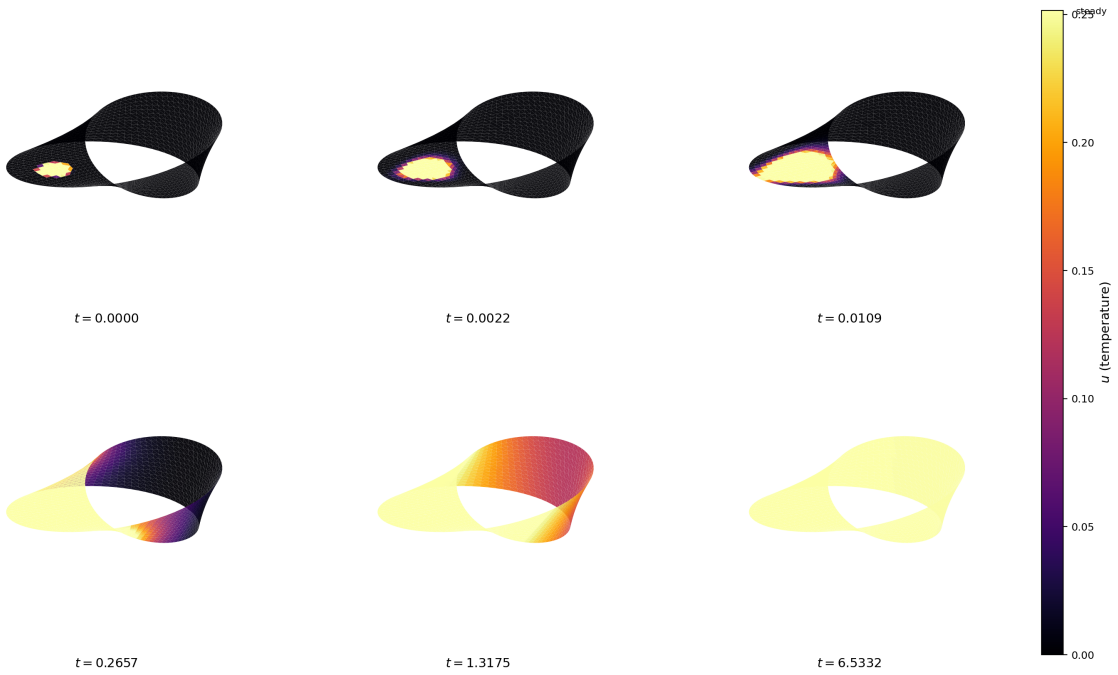


Figure 5: Heat diffusion on the Möbius strip at six logarithmically spaced times, solved via implicit Euler time-stepping with the discrete Laplace-Beltrami operator under homogeneous Neumann boundary conditions. The temperature field u spreads from an initial Gaussian spike and converges toward a uniform steady state $\bar{u} =$ total heat/total area. The non-orientability of the strip is visible in the asymmetric spreading of the wavefront across the seam.

The Galerkin method can be understood as projecting a differential operator onto a high-order function basis, such as a monomial or polynomial basis. See [48] for an overview. To start, we will translate the Laplacian eigenvalue problem to a variational problem using Green's formula—that is

$$\int_{\mathbb{M}} \varphi \Delta_{\mathbb{M}} f \, d\sigma = -\int_{\mathbb{M}} \langle \nabla_{\mathbb{M}} f, \nabla_{\mathbb{M}} \varphi \rangle \, d\sigma, \tag{5}$$

where $d\sigma$ is a surface or volume element. Note that we are assuming the line integral term vanishes, so it is necessary to assume that either $\mathbb{M}$ is without boundary or that we have Dirichlet or Neumann boundary conditions. Note that $\varphi \in L^2(\mathbb{M})$ is a test function, and by multiplying the Helmholtz operator by φ we may use Equation 5 to obtain

$$\varphi \Delta_{\mathbb{M}} f = -\lambda \varphi f \iff \int_{\mathbb{M}} \varphi \Delta_{\mathbb{M}} f \, d\sigma = -\lambda \int_{\mathbb{M}} \varphi f \, d\sigma \iff \int_{\mathbb{M}} \langle \nabla_{\mathbb{M}} f, \nabla_{\mathbb{M}} \varphi \rangle \, d\sigma = \lambda \int_{\mathbb{M}} \varphi f \, d\sigma.$$

To proceed, use the definition of $\nabla_{\mathbb{M}}$ in local coordinates so that

$$\begin{aligned}
\langle \nabla_{\mathbb{M}} f, \nabla_{\mathbb{M}} \varphi \rangle &= \left\langle \sum_{i,j} g^{ij} \partial_i f \partial_j, \sum_{k,l} g^{kl} \partial_k \varphi \partial_l \right\rangle \\
&= \sum_{i,j} g^{ij} f \partial_j \left(\sum_{k,l} g^{kl} \partial_l \varphi \left\langle \partial_i, \partial_k \right\rangle \right) \\
&= \sum_{i,j} g^{ij} f \partial_j \varphi \left(\sum_{k,l} g^{kl} g_{ik} \partial_l \right) \\
&= \sum_{i,j} g^{ij} f \partial_j \varphi \left(\sum_{k,l} \delta_{ki} \delta_{li} \partial_l \right) \\
&= \sum_{i,j} g^{ij} \partial_i f \partial_j \varphi.
\end{aligned}$$

Taking this with our above computation, it follows that

$$\int_{\mathbb{M}} \sum_{i,j} g^{ij} \partial_i f \partial_j \varphi \, d\sigma = \lambda \int_{\mathbb{M}} \varphi f \, d\sigma. \tag{6}$$

Now we transform this into a generalized eigenvalue problem. Choose a finite set of linearly independent *trial functions* of the form $F_1, \ldots, F_N : \mathbb{M}' \to \mathbb{R}$ where $\mathbb{M}'$ is a parameterization of $\mathbb{M}$ by a mesh. We want to approximate solutions of the form $f = \sum_k c_k F_k$ for unknown $c_k \in \mathbb{R}$. Thus, we substitute this choice of f into Equation 6 and choose a family of test functions to be precisely the family of trial functions $\{F_l\}$ for $l = 1, \ldots, N$ to obtain a family of *Galerkin equations* of the form

$$\int_{\mathbb{M}} \sum_{i,j} g^{ij} \partial_i \left(\sum_k c_k F_k \right) \partial_j F_l \, d\sigma = \lambda \int_{\mathbb{M}} F_l f \, d\sigma. \tag{7}$$

Now define matrices $A, B \in \mathbb{R}^{N \times N}$ such that

$$A_{kl} = \int_{\mathbb{M}} \sum_{i,j} g^{ij} \partial_i F_k \partial_j F_l \, d\sigma \quad \text{and} \quad B_{kl} = \int_{\mathbb{M}} F_l F_k \, d\sigma.$$

By our choice of test functions we see that A and B are symmetric. Furthermore, these matrices are positive semi-definite, just as the operator that we discretized to construct them. Finally, observe that taking a unknown coefficient vector $\mathbf{c} \in \mathbb{R}^N$, then the generalized eigenproblem $A\mathbf{c} = \lambda B\mathbf{c}$ is precisely the family of Galerkin equations defined by Equation 7. As noted in [42], this method is essentially independent of a choice of mesh.

5.4 Laplacian Autoencoders

Another application of the Laplace-Beltrami operator relevant to classwork study is the Laplacian autoencoder, where an autoencoder is trained completely from a manifold learning framework, explored in [28]. It regularizes autoencoder training so that local linearity is preserved, just as the LBO (and graph Laplacian) does. The LBO specifically allows for discrete approximations of first- and higher-order regularizers. The authors describe a general objective, $\mathcal{J}(f, g) = \varepsilon(f, g) + \lambda\Omega(f)$, where f is the encoder, g is the decoder, ε the error term, and Ω the regularization term. They found that for a higher-derivative case, the optimal choice of $\Omega(f)$ was $\int_{\mathbb{M}} \|\nabla f(x)\|_2^2 dx + \int_{\mathbb{M}} \|H_f(x)\|_F^2 dx$ when attempting to preserve locality, where H_f is the Hessian of f. In the first-order case, the optimal regularization term is just $\int_{\mathbb{M}} \|\nabla f(x)\|_2^2 dx$. From there, $\int_{\mathbb{M}} \|\nabla f(x)\|_2^2 dx$ in the discrete case of the weight matrix between data points becomes $\sum_{i<j} W(i,j) \left(f(x_i) - f(x_j)\right)^2$ (as shown in [4]) and [28] demonstrate using similar techniques that the discretization of $\int_{\mathbb{M}} \|H_f(x)\|_F^2 dx$ is $\sum_{i<j} W(i,j) \left(\nabla f(x_i) - \nabla f(x_j)\right)^2$. These discretizations can be used to find a practical training objective, either

$$\mathcal{J}(f, g) = \sum_{x \in S} E(x, g(f(x)) + \lambda \frac{1}{2} \sum_{x_i, x_j \in S} W(i,j) \|f(x_i) - f(x_j)\|_2^2$$

for a first-order Laplacian autoencoder, or

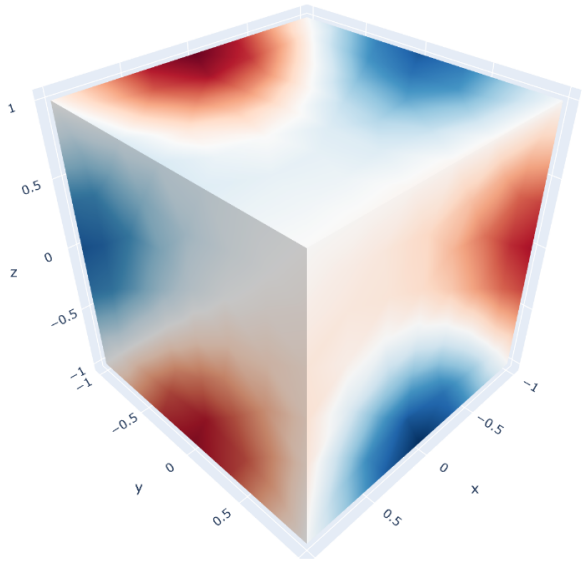


Figure 6: A mesh parameterization for the surface of a cube is colored with intensity determined by the computed coefficients $\mathbf{c}$ from the generalized eigenvalue problem $A\mathbf{c} = \lambda B\mathbf{c}$. Specifically, these coefficients come from the fifteenth eigenfunction of the Laplace-Beltrami operator. Notice that the heat map induces contours that partition the surface.

$$\mathcal{J}(f,g) = \sum_{x \in S} E(x, g(f(x)) + \frac{1}{2} \sum_{x_i, x_j \in S} W(i,j) \left(\lambda \|f(x_i) - f(x_j)\|_2^2 + \beta \|J_f(x_i) - J_f(x_j)\|_2^2\right)$$

for a higher-dimensional LAE. Here, E is again an error term, chosen based on the data (but usually either squared loss or cross-entropy loss), and the reconstruction function is chosen based on that function. Back propagation is used to determine the gradient of $\mathcal{J}$, and the authors found that the Laplacian autoencoder outperformed similiar methods, perhaps because it relies on the innate manifold the data lie on.

6 Conclusion

The Laplace-Beltrami Operator, while coming from the definition of the Laplacian in calculus, has far-reaching application when generalized to a compact Riemannian manifold. We have shown how its properties, specifically preservation of locality, self-adjointness, and positive semidefiniteness, bring meaning to many methods in dimensionality reduction, as eigenvalues and vectors gain meaningful order and can be used as tools to model aspects of the manifold. Additionally, preservation of locality shows that the Laplace-Beltrami operator is a useful transformation when the goal is to keep close points together, but also gives new validation to the methods of Laplacian eigenmaps and Local Linear Embedding, both of which rely on that preservation to accurately embed data on a manifold into a lower-dimensional space. By studying the Laplace-Beltrami operator, one is able to understand the innate properties that allow methods of dimensionality reduction to preserve local structure, and extend this paradigm to a broader domain of problems and applications.

References

[1] Armen Aghajanyan, Sonal Gupta, and Luke Zettlemoyer. Intrinsic dimensionality explains the effectiveness of language model fine-tuning. In *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: long papers)*, pages 7319–7328, 2021.

[2] Yulin An and Enrique del Castillo. An ai approach for learning the spectrum of the laplace-beltrami operator. *arXiv preprint arXiv:2507.07073*, 2025.

[3] Farzana Anowar, Samira Sadaoui, and Bassant Selim. Conceptual and empirical comparison of dimensionality reduction algorithms (pca, kpca, lda, mds, svd, lle, isomap, le, ica, t-sne). *Computer Science Review*, 40:100378, 2021.

[4] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, 2003.

[5] Mikhail Belkin and Partha Niyogi. Convergence of laplacian eigenmaps. In *Advances in Neural Information Processing Systems (NIPS)*, volume 19, page 129. MIT Press, 2006.

[6] Mikhail Belkin and Partha Niyogi. Towards a theoretical foundation for laplacian-based manifold methods. *Journal of Computer and System Sciences*, 74(8):1289–1308, 2008.

[7] Mikhail Belkin, Jian Sun, and Yusu Wang. Discrete laplace operator on meshed surfaces. In *Proceedings of the twenty-fourth annual symposium on Computational geometry*, pages 278–287, 2008.

[8] Håvard Bakke Bjerkevik, Magnus Bakke Botnan, and Michael Kerber. Computing the interleaving distance is np-hard. *Foundations of Computational Mathematics*, 20(5):1237–1271, 2020.

[9] Mario Botsch, Leif Kobbelt, Mark Pauly, Pierre Alliez, and Bruno Lévy. *Polygon Mesh Processing*. AK Peters, Natick, Massachusetts, 2010.

[10] David Cohen-Steiner, Herbert Edelsbrunner, and John Harer. Stability of persistence diagrams. In *Proceedings of the twenty-first annual symposium on Computational geometry*, pages 263–271, 2005.

[11] David Cohen-Steiner, Herbert Edelsbrunner, and Dmitriy Morozov. Vines and vineyards by updating persistence in linear time. In *Proceedings of the twenty-second annual symposium on Computational geometry*, pages 119–126, 2006.

[12] Richard Courant and David Hilbert. *Methods of mathematical physics, volume 2*. John Wiley & Sons, 2024.

[13] Michael AA Cox and Trevor F Cox. Multidimensional scaling. In *Handbook of data visualization*, pages 315–347. Springer, 2008.

[14] Feng Dai and Yuan Xu. Springer Monographs in Mathematics. Springer New York, NY, 2013. doi:10.1007/978-1-4614-6660-4.

[15] Vin De Silva and Mikael Vejdemo-Johansson. Persistent cohomology and circular coordinates. In *Proceedings of the twenty-fifth annual symposium on Computational geometry*, pages 227–236, 2009.

[16] Mathieu Desbrun. Applied geometry: Discrete differential calculus for graphics. In *Computer Graphics Forum*, volume 23, pages 269–269. Wiley Online Library, 2004.

[17] Shen Dong, Peer-Timo Bremer, Michael Garland, Valerio Pascucci, and John C Hart. Quadrangulating a mesh using laplacian eigenvectors. 2005.

[18] George H Dunteman. *Principal components analysis*, volume 69. Sage, 1989.

[19] Edelsbrunner, Letscher, and Zomorodian. Topological persistence and simplification. *Discrete & computational geometry*, 28(4):511–533, 2002.

[20] Charles Fefferman, Sanjoy Mitter, and Hariharan Narayanan. Testing the manifold hypothesis. *Journal of the American Mathematical Society*, 29(4):983–1049, 2016.

[21] Paul Finsler. *Über kurven und flächen in allgemeinen räumen*. Springer-Verlag, 2013.

[22] Peter Gabriel. Unzerlegbare darstellungen i. *Manuscripta mathematica*, 6(1):71–103, 1972.

[23] Claudia Garcia, Zenab Hassainia, and Emeric Roulley. Dynamics of vortex cap solutions on the rotating unit sphere. *arXiv preprint arXiv:2306.00154*, 2023. URL: https://arxiv.org/abs/2306.00154.

[24] Michael C Hout, Megan H Papesh, and Stephen D Goldinger. Multidimensional scaling. *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(1):93–103, 2013.

[25] Alan Julian Izenman. Introduction to manifold learning. *Wiley Interdisciplinary Reviews: Computational Statistics*, 4(5):439–446, 2012.

[26] Dmitry Jakobson, Nikolai Nadirashvili, and John Toth. Geometric properties of eigenfunctions. *Russian Mathematical Surveys*, 56(6):1085–1105, 2001.

[27] Layla Jarrahy, Christian Lentz, and Vinith Yedidi. Laplace beltrami project - jarrahy, lentz, yedidi. URL: https://github.com/ChristianLentz/Laplace-Beltrami-Project-Jarrahy-Lentz-Yedidi/tree/main.

[28] Kui Jia, Lin Sun, Shenghua Gao, Zhan Song, and Bertram Shi. Laplacian auto-encoders: An explicit learning of nonlinear data manifold. *Neurocomputing*, 2015. URL: http://dx.doi.org/10.1016/j.neucom.2015.02.023.

[29] Priyanka Jindal and Dharmender Kumar. A review on dimensionality reduction techniques. *Int. J. Comput. Appl*, 173(2):42–46, 2017.

[30] Joseph B Kruskal and Myron Wish. *Multidimensional scaling*. Number 11. Sage, 1978.

[31] Michael Lesnick. The theory of the interleaving distance on multidimensional persistence modules. *Foundations of Computational Mathematics*, 15(3):613–650, 2015.

[32] Bruno Lévy. Laplace-beltrami eigenfunctions towards an algorithm that 'understands' geometry. In *IEEE International Conference on Shape Modeling and Applications 2006 (SMI'06)*, pages 13–13. IEEE, 2006.

[33] Yunqian Ma and Yun Fu. *Manifold learning theory and applications*, volume 434. CRC press Boca Raton, 2012.

[34] Marina Meilă and Hanyu Zhang. Manifold learning: What, how, and why. *Annual Review of Statistics and Its Application*, 11(1):393–417, 2024.

[35] John Nash. The imbedding problem for riemannian manifolds. *Annals of mathematics*, 63(1):20–63, 1956.

[36] Ahmad Nasikun, Christopher Brandt, and Klaus Hildebrandt. Fast approximation of laplace-beltrami eigenproblems. In *Computer Graphics Forum*, volume 37, pages 121–134. Wiley Online Library, 2018.

[37] Peter J. Olver et al. *Introduction to partial differential equations*, volume 1. Springer, 2014.

[38] Ulrich Pinkall and Konrad Polthier. Computing discrete minimal surfaces and their conjugates. *Experimental mathematics*, 2(1):15–36, 1993.

[39] John William Strutt Baron Rayleigh. *The theory of sound*, volume 2. Macmillan, 1896.

[40] G Thippa Reddy, M Praveen Kumar Reddy, Kuruva Lakshmanna, Rajesh Kaluri, Dharmendra Singh Rajput, Gautam Srivastava, and Thar Baker. Analysis of dimensionality reduction techniques on big data. *Ieee Access*, 8:54776–54788, 2020.

[41] Martin Reuter, Silvia Biasotti, Daniela Giorgi, Giuseppe Patanè, and Michela Spagnuolo. Discrete laplace–beltrami operators for shape analysis and segmentation. *Computers & Graphics*, 33(3):381–390, 2009.

[42] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace–beltrami spectra as 'shape-dna'of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.

[43] Steven Rosenberg. *The Laplacian on a Riemannian manifold: an introduction to analysis on manifolds*. Number 31. Cambridge University Press, 1997.

[44] Sam T Roweis and Lawrence K Saul. Nonlinear dimensionality reduction by locally linear embedding. *science*, 290(5500):2323–2326, 2000.

[45] Bernhard Schölkopf, Alexander Smola, and Klaus-Robert Müller. Kernel principal component analysis. In *International conference on artificial neural networks*, pages 583–588. Springer, 1997.

[46] Lindsay I Smith. A tutorial on principal components analysis. 2002.

[47] Vincent Spruyt. The curse of dimensionality in classification. *Comput. Vis. Dummies*, 21(3):35–40, 2014.

[48] Gilbert Strang et al. *Introduction to applied mathematics*, volume 16. Wellesley-Cambridge Press Wellesley, MA, 1986.

[49] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28, pages 1383–1392. Wiley Online Library, 2009.

[50] Joshua B Tenenbaum, Vin de Silva, and John C Langford. A global geometric framework for nonlinear dimensionality reduction. *science*, 290(5500):2319–2323, 2000.

[51] Vidar Thomée. *Galerkin Finite Element Methods for Parabolic Problems*, volume 25 of *Springer Series in Computational Mathematics*. Springer, Berlin, 2 edition, 2006.

[52] Loring W Tu. *An Introduction to Manifolds*. Springer, 2011.

[53] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[54] Laurens Van Der Maaten, Eric O Postma, H Jaap Van Den Herik, et al. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research*, 10(1):1–41, 2009.

[55] Libor Váša, Stefano Marras, Kai Hormann, and Guido Brunnett. Compressing dynamic meshes with geometric laplacians. In *Computer Graphics Forum*, volume 33, pages 145–154. Wiley Online Library, 2014.

[56] Michel Verleysen and Damien François. The curse of dimensionality in data mining and time series prediction. In *International work-conference on artificial neural networks*, pages 758–770. Springer, 2005.

[57] Max Wardetzky. Convergence of the cotangent formula: An overview. *Discrete differential geometry*, pages 275–286, 2008.

[58] Simon Weber, Thomas Dages, Maolin Gao, and Daniel Cremers. Finsler-laplace-beltrami operators with application to shape analysis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 3131–3140, 2024.

[59] Chenkai Xu, Hongwei Lin, Hui Hu, and Yaqi He. Fast calculation of laplace-beltrami eigenproblems via subdivision linear subspace. *Computers & graphics*, 97:236–247, 2021.

[60] Guoliang Xu. Discrete laplace–beltrami operators and their convergence. *Computer aided geometric design*, 21(8):767–784, 2004.

[61] Ao Zhang, Qing Fang, Peng Zhou, and Xiao-Ming Fu. Topology-controlled laplace–beltrami operator on point clouds based on persistent homology. *Graphical Models*, 139:101261, 2025.

[62] Afra Zomorodian and Gunnar Carlsson. Computing persistent homology. In *Proceedings of the twentieth annual symposium on Computational geometry*, pages 347–356, 2004.