

From Varadhan’s Limit to Eigenmaps: A Guide to the Geometric Analysis behind Manifold Learning

Chen-Yun Lin ¹ and Christina Sormani ²

¹Lehman College, 250 Bedford Park Boulevard West, Bronx, NY 10468, USA

²CUNY Graduate Center, 365 Fifth Avenue, NY, NY 10468, USA

*Dedicated to Professor Varadhan,
Fellow of The World Academy of Sciences*

Abstract. We present an overview of the history of the heat kernel and eigenfunctions on Riemannian manifolds and how the theory has led to modern methods of analyzing high dimensional data.

Keywords: heat kernel embeddings, eigenmaps, manifold learning.

2020 Mathematics Subject Classification: 58J35, 35K05, 53C23.

Contents

1	Introduction	34
2	Heat Kernels and Distances on Riemannian Manifolds	35
2.1	The Heat Kernel on a Circle	36
2.2	The Heat Kernel on a Rescaled Circle	36
2.3	The Heat Kernel on Tori	37
2.4	The Heat Kernel on Spaces of Constant Curvature	37
2.5	Estimating the Heat Kernel on Families of Manifolds	38

¹Partially supported by PSC-CUNY Grant. E-mail: chenyunlin.math@gmail.com

²Partially supported by partially supported by NSF DMS 1612049 and PSC-CUNY Grants.
E-mail: sormanica@gmail.com

3	Eigenfunctions and Eigenvalues	40
3.1	Finding Eigenfunctions using the Rayleigh Quotient	41
3.2	The Sturm-Liouville Decomposition of the Heat Kernel	42
3.3	Truncating the Sturm-Liouville Expansion of the Heat Kernel	43
3.4	Eigenvalues and Eigenfunctions on a Standard Circle	44
3.5	Eigenvalues and Eigenfunctions on a Rescaled Circle	45
3.6	Eigenvalues and Eigenfunctions on Tori	46
3.7	Eigenvalues and Eigenfunctions on a Thin Torus	48
3.8	Eigenvalues and Eigenfunctions on a Sphere	49
3.9	Estimates on the First Eigenvalue	50
3.10	Estimating Eigenvalues on Families of Manifolds	51
4	Classes of Riemannian Manifolds and Convergence	51
4.1	Smooth, C^0 , $C^{k,\alpha}$, and C^k Convergence	52
4.2	Smooth Convergence of Tori and Warped Tori	54
4.3	Convergence of Distances between Points	55
4.4	Lipschitz Convergence	55
4.5	C^2 convergence preserves Δ and $h_t(x, y)$	56
4.6	Diffeomorphic Compactness Theorems	57
4.7	Injectivity Radius, Heat Kernels, and Eigenvalues	58
4.8	Harmonic Coordinates and Harmonic Radius	59
4.9	Gromov-Hausdorff Convergence	59
4.10	Increasingly Thin Tori GH-Converge to Circles	61
4.11	Eigenvalues and Measured Gromov-Hausdorff Convergence	62
4.12	Barbells	64
4.13	Intrinsic Flat Convergence	66
4.14	VADB Convergence	68
5	Embedding Riemannian Manifolds via the Heat Kernel	68
5.1	Rescaled Heat Kernel Embeddings	70
5.2	Truncated Heat Kernel Embeddings	71
5.3	Embedding Circles with a Truncated Heat Kernel	72
5.4	Embedding Spheres with a Truncated Heat Kernel	74
5.5	Embedding Tori with a Truncated Heat Kernel Map	75
5.6	Embedding Tori with Selected Eigenfunctions	75
5.7	Estimating Lengths vs Extrinsic Distances	77
5.8	Other Rescalings of the Heat Kernel Embedding	79
5.9	Spectral Convergence of Riemannian Manifolds	79

6	Applications to Dimension Reduction of High Dimensional Data Sets	79
6.1	Belkin-Niyogi Eigenmaps of Data Sets	82
6.2	A Brief History of Graph Laplacians starting with Dodziuk . . .	84
6.3	Continuing the History of Graph Laplacians with work of Chung	85
6.4	Graph Laplacians Approximating Smooth Laplacians	86
6.5	An Eigenmap of a Hexagon	86
6.6	Finding the Eigenmap of a Data Set of 3×3 Photos	89
6.7	An Eigenmap of a Sphere	90
6.8	Eigenmaps of Barbells	92
6.9	Eigenmaps of Lollipops	93
6.10	Coifman and Lafon's Diffusion Maps of Data Sets	95
6.11	Diffusion Distance	96
7	Spectral Embeddings and Embedding Dimensions	97
7.1	Deforming Manifolds and their Embeddings	98
7.2	Local Embeddings of M^n into \mathbb{E}^n	99
7.3	Embedding Dimensions of Spectral Embeddings	99
7.4	Uniform Bounds on Spectral Embeddings by Portegies	100
7.5	Spectral Embeddings depending on the Connection Laplacian .	101
7.6	L^p Bounds on Pull Back Metrics	102
8	Open Questions	103
8.1	Conjectures on Truncated Eigenmaps of Small Dimension . . .	103
8.2	Conjectures on Selective Eigenmaps	104
8.3	Conjectures on Convergence of Pullbacks of Embeddings	105
8.4	Conjectures on Convergence of Manifolds and their Embeddings	106

1 Introduction

Geometry is one of the most beautiful and ancient forms of mathematics. Inspired by Professor Romain Murenzi's speech at the *2022 TWAS TYAN Virtual Workshop on Differential Geometry*, we will present the Geometric Analysis that leads a very young and active branch of Geometry called Manifold Learning. Manifold Learning is a form of machine learning that has applications in biology, computer vision, natural language processing, and data mining. It is built upon half a century of Geometric Analysis. Here we review the history from Varadhan's theorem relating the heat kernel and the distance on a Riemannian manifold, to the study of eigenfunctions and eigenvalues in relation to the

heat kernel, to convergence of Riemannian manifolds, to embeddings of Riemannian manifolds using the heat kernel, to the modern notions of eigenmaps and dimension reduction through manifold learning. Throughout we provide detailed examples demonstrating the ideas. We assume only knowledge of Professor Manfredo P. do Carmo's **Riemannian Geometry** and we provide a guide to both the original papers and various useful textbooks.

2 Heat Kernels and Distances on Riemannian Manifolds

We will begin with a 1967 theorem of TWAS Fellow and Abel Laureate, Srinivasan Varadhan, who earned his doctorate at the Indian Statistical Institute in 1963:

Theorem 2.1. [93] The distance between two points, x and y in a compact Riemannian manifold, M , satisfies

$$d(x, y)^2 = \lim_{t \rightarrow 0} -4t \log h_t(x, y) \quad (2.1)$$

where $h_t(x, y)$ is the heat kernel. This convergence is uniform on bounded regions, U , in the sense that: $\forall \varepsilon > 0 \exists T_{U, \varepsilon} > 0$ such that $\forall t \in (0, T_{U, \varepsilon})$ we have

$$|d(x, y)^2 + 4t \log(h_t(x, y))| < \varepsilon \quad (2.2)$$

which is equivalent to saying that

$$e^{-d^2(x, y)/(4t)} \cdot e^{-\varepsilon/(4t)} < h_t(x, y) < e^{-d^2(x, y)/(4t)} \cdot e^{\varepsilon/(4t)}. \quad (2.3)$$

Recall that the heat kernel

$$h_t : M \times M \rightarrow (0, \infty) \text{ defined for } t > 0 \quad (2.4)$$

is defined to be the unique function such that

$$u(x, t) = \int_M f(y) h_t(x, y) d\text{vol}_y, \quad (2.5)$$

for any solution, $u : M \times (0, \infty) \rightarrow [0, \infty)$ of the heat equation,

$$\Delta_x u(x, t) = \partial_t u(x, t) \text{ with initial data } \lim_{t \rightarrow 0^+} u(x, t) = f(x). \quad (2.6)$$

As a consequence, for any $x \in M$,

$$\int_M h_t(x, y) d\text{vol}_y = 1. \quad (2.7)$$

Furthermore, it is easy to verify the heat semigroup property holds:

$$\int_M h_t(x, y) h_s(y, z) d\text{vol}_y = h_{t+s}(x, z). \quad (2.8)$$

For a good resource with the details as to why the heat kernel exists on a compact Riemannian manifold and proofs of its various properties, we refer the reader to Chavel's textbook **Eigenvalues in Riemannian Geometry** [22]. For more theorems about the heat kernel see Schoen-Yau's text **Lectures on Differential Geometry** [83].

Varadhan proved his theorem by estimating the heat kernel on normal neighborhoods using the Euclidean heat kernel and then built his estimate on $h_t(x, y)$ for x far from y by applying the heat semigroup property as well as the properties of the exponential map based at x near the minimizing geodesic from x to y . This proof appeared in [93] and is quite readable for geometers. He wrote a probabilistic proof in [92].

2.1 The Heat Kernel on a Circle

On a circle, \mathbb{S}^1 parametrized by $\theta \in [0, 2\pi)$ with $g_{\mathbb{S}^1} = d\theta^2$, then the Laplacian is $\Delta = \partial_\theta^2$ and the heat kernel is the following sum which converges due to the exponential decay:

$$h_t(\theta_1, \theta_2) = \sum_{m \in \mathbb{Z}} \frac{1}{\sqrt{4\pi t}} e^{-(\theta_1 - \theta_2 + 2\pi m)^2 / (4t)}. \quad (2.9)$$

Notice that as $t \rightarrow 0$, the term in the sum which dominates because it decays the least quickly, is the term with

$$e^{-(\theta_1 - \theta_2 + 2\pi m)^2 / (4t)} = e^{-d_{\mathbb{S}^1}^2(\theta_1, \theta_2)^2 / (4t)} \quad (2.10)$$

where $d_{\mathbb{S}^1}(\theta_1, \theta_2)$ is the Riemannian distance between the points in \mathbb{S}^1 :

$$d_{\mathbb{S}^1}(\theta_1, \theta_2) = \min\{|\theta_1 - \theta_2 + 2\pi m| : m \in \mathbb{Z}\} \quad (2.11)$$

which is the length of the smaller arc between θ_1 and θ_2 .

2.2 The Heat Kernel on a Rescaled Circle

On a circle, \mathbb{S}_R^1 , we rescale the circle by a factor $R > 0$ so that $x = R\theta \in [0, 2\pi R)$ with $g_R = dx^2 = R^2 d\theta^2$. Then the Laplacian rescales $\Delta_R = \partial_x^2 = R^{-2} \partial_\theta^2$ and the heat kernel is

$$h_t^R(\theta_1, \theta_2) = \sum_{m \in \mathbb{Z}} \frac{1}{\sqrt{4\pi t}} e^{-(\theta_1 - \theta_2 + 2\pi m)^2 R^2 / (4t)}. \quad (2.12)$$

Note that the heat kernel of \mathbb{S}_R^1 can be related to the heat kernel of \mathbb{S}^1 with dilated time:

$$h_t^R(\theta_1, \theta_2) = R^{-1} h_{t/R^2}(\theta_1, \theta_2) \quad (2.13)$$

because

$$\sum_{m \in \mathbb{Z}} \frac{1}{\sqrt{4\pi t}} e^{-(\theta_1 - \theta_2 + 2\pi m)^2 R^2 / (4t)} = R^{-1} \sum_{m \in \mathbb{Z}} \frac{1}{\sqrt{4\pi(t/R^2)}} e^{-(\theta_1 - \theta_2 + 2\pi m)^2 / (4(t/R^2))} \quad (2.14)$$

2.3 The Heat Kernel on Tori

Next let us consider the flat torus formed by taking the isometric product of rescaled circles,

$$\mathbb{S}_A^1 \times \mathbb{S}_B^1 = (\mathbb{S}^1 \times \mathbb{S}^1, A^2 d\theta^2 + B^2 d\varphi^2). \quad (2.15)$$

Its heat kernel is

$$h_t^{A,B}((\theta_1, \varphi_1), (\theta_2, \varphi_2)) = \sum_{n, m \in \mathbb{Z}} \frac{1}{(4\pi t)} e^{-((\theta_1 - \theta_2 + 2\pi n)^2 A^2 + (\varphi_1 - \varphi_2 + 2\pi m)^2 B^2) / (4t)}. \quad (2.16)$$

Notice that

$$4t \log h_t^{A,B}((\theta_1, \varphi_1), (\theta_2, \varphi_2)) = 4t \log h_t^A(\theta_1, \theta_2) + 4t \log h_t^B(\varphi_1, \varphi_2). \quad (2.17)$$

Since the distances squared also add, we have

$$d_{\mathbb{S}_A^1 \times \mathbb{S}_B^1}^2((\theta_1, \varphi_1), (\theta_2, \varphi_2)) = d_{\mathbb{S}_A^1}^2(\theta_1, \theta_2) + d_{\mathbb{S}_B^1}^2(\varphi_1, \varphi_2) \quad (2.18)$$

which matches what Varadhan's limit predicts in this setting.

2.4 The Heat Kernel on Spaces of Constant Curvature

In general on a homogeneous manifold, M^n , the heat kernel will be a function only of distance and time:

$$h_t^M(x, y) = h_t^M(s) \text{ where } s = d(x, y). \quad (2.19)$$

The heat kernel on a sphere was first computed by Minakshisundaram [76]. The heat kernel on hyperbolic space was first computed in dimension 2 to be

$$h_t^{\mathbb{H}^2}(s) = \frac{\sqrt{2}}{(4\pi t)^{3/2}} e^{-t/4} \int_s^\infty \frac{u e^{-u^2/(4t)}}{(\cosh(u) - \cosh(s))^{1/2}} du \quad (2.20)$$

by McKean [74]. In three dimensions it was shown to be

$$h_t^{\mathbb{H}^3}(s) = \frac{(-1)}{2\pi} \frac{1}{\sqrt{4\pi t}} \left(\frac{1}{\sinh(s)} \frac{\partial}{\partial s} \right) e^{-t} e^{-s^2/4t} \quad (2.21)$$

by Davies-Mandouvalos in [41]. They applied an unpublished recurrence relation proved by Millson:

$$h_t^{\mathbb{H}^{m+2}}(s) = -\frac{e^{-mt}}{2\pi \sinh(s)} \frac{\partial}{\partial s} h_t^{\mathbb{H}^m}(s) \quad (2.22)$$

to provide the formula for the heat kernel in all dimensions. See also the alternative proof by Grigor'yan-Noguchi in [54].

When M_κ^n is the simply connected Riemannian manifold of constant curvature κ , we write

$$h_t^{k,n}(s) = h_t^{M_\kappa^n}(s). \quad (2.23)$$

Since M_κ^n is just a rescaled sphere $\mathbb{S}_{\kappa^{-1/2}}^n$ when $\kappa > 0$ and is just rescaled hyperbolic space $\mathbb{H}_{\kappa^{-1/2}}^n$ when $\kappa < 0$, we know

$$h_t^{k,n}(s) = \kappa^{n/2} h_{\kappa t}^{\mathbb{S}^n}(\kappa^{1/2}s) \text{ for } \kappa > 0 \quad (2.24)$$

and

$$h_t^{k,n}(s) = |\kappa|^{n/2} h_{|\kappa|t}^{\mathbb{H}^n}(|\kappa|^{1/2}s) \text{ for } \kappa < 0. \quad (2.25)$$

2.5 Estimating the Heat Kernel on Families of Manifolds

In 1976 Debiard-Gaveau-Mazet [42] estimated the heat kernel uniformly from above and below.

Theorem 2.2. For simply connected manifolds, M , with negative sectional curvature between bounds

$$\kappa \leq \text{sect}_M \leq \kappa' < 0, \quad (2.26)$$

the heat kernel satisfies

$$h_t^{\kappa',n}(s) \geq h_t^M(x,y) \geq h_t^{\kappa,n}(s). \quad (2.27)$$

where $h^{\kappa,n}(t)(s)$ is defined as in (2.23) and where $s = d_M(x,y)$.

Applying this theorem, we thus have

$$s^2 + 4t \log h_t^{\kappa,n}(s) \geq d_M(x,y)^2 + 4t \log h_t^M(x,y) \geq s^2 + 4t \log h_t^{\kappa',n}(s). \quad (2.28)$$

So $|d_M(x, y)^2 + 4t \log h_t^M(x, y)|$ is bounded above by

$$\max \left\{ |s^2 + 4t \log h_t^{K,n}(s)|, |s^2 + 4t \log h_t^{K',n}(s)| \right\}. \quad (2.29)$$

Thus the uniform convergence in Varadhan's Theorem is uniform even allowing the Riemannian manifold to vary within this class of Riemannian manifolds M such that (2.26) holds.

In 1981, Cheeger and Yau [27] proved the following estimate that only requires uniform lower bounds on the Ricci curvature.

Theorem 2.3. If M^n has $\text{Ricci} \geq (n - 1)H$ for any $H \in (-\infty, \infty)$, then we still have a lower bound

$$h_t^M(x, y) \geq h_t^{H,n}(d_M(x, y)). \quad (2.30)$$

So if M is also simply connected with $\text{sect}_M \leq \kappa'$ we still have (2.27) and its consequences.

In 1982, Cheeger-Gromov-Taylor found an upper bound on the heat kernel for points $x, y \in B(p, D)$ depending on D , on the injectivity radius and the volumes of balls in the manifold (See Theorem 1.4 in [26]). In 1984 Chavel found an upper bound on the heat kernel which depends on isoperimetric constants of balls (See Theorems 8 and 9 in Chapter VIII of [22]) thus there is a uniform bound on the Varadhan time for manifolds with same isoperimetric profile functions and the same lower bound on Ricci curvature. These bounds may also be used to provide uniform control on the heat kernel.

Cheng-Li-Yau studied complete manifolds and discussed some of the difficulties proving upper bounds before finding an upper bound on the growth of the heat kernel [30]. Cheng and Li found upper bounds on the heat kernel depending on the Sobolev constant of the manifold [29]. Additional bounds on the heat kernel and also a Harnack Inequality for Riemannian manifolds is proven by Li and Yau in 1986 [67]. This was then applied by Li in [66] to estimate the behavior of the heat kernel as $t \rightarrow \infty$. For more details see the text of Schoen and Yau [83].

It should be noted that Varadhan's original proof of his limit in [92] was completed using probabilistic methods and there has been much work done using these methods that we do not present here. These results would require an expertise in probability and would better be presented elsewhere.

3 Eigenfunctions and Eigenvalues

One of the most beautiful properties of the heat kernel on a compact Riemannian manifold is its Sturm-Liouville decomposition:

$$h_t(x, y) = \sum_{j=0}^{\infty} e^{-\lambda_j t} \phi_j(x) \phi_j(y) \quad (3.1)$$

where $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$ are eigenvalues and ϕ_j are an L^2 orthonormal collection of eigenfunctions of the Laplacian:

$$\Delta \phi_j(x) = -\lambda_j \phi_j(x) \quad (3.2)$$

For a quick hint at why this expansion gives the heat kernel, note that each term satisfies the heat equation:

$$\partial_t e^{-\lambda_j t} \phi_j(x) = -\lambda_j e^{-\lambda_j t} \phi_j(x) = \Delta e^{-\lambda_j t} \phi_j(x). \quad (3.3)$$

It is also symmetric $h_t(x, y) = h_t(y, x)$. We will explain in more detail below.

Note that Sturm-Liouville theory is named for Jacques Charles Francois Sturm (1803-1855) and Joseph Liouville (1809-1882). From the age of 16, Sturm supported his mother and siblings as a mathematics tutor: he moved to Paris to serve as a tutor to a student of his that continued on to study there. He became a secondary school teacher after the Revolution of 1830 and by 1834 had published his work developing Sturm-Liouville Theory with Joseph Liouville. Joseph Liouville was the son of an army officer in France and had studied engineering at Ecole Polytechnique, but was working as a mathematics assistant at various institutions at the time. After the publication of this important work and other theorems, both became professors of mathematics at Ecole Polytechnique in 1838.

See also work of Miranda [78] applying Sturm-Liouville Theory to Riemannian manifolds. This is well explained in Chavel's textbook [22] and Schoen-Yau's text [83]. For a survey of earlier results, the text of Berger, P. Gauduchon and E. Mazet, [17] is often recommended. This last book includes many explicit computations of eigenvalues and eigenfunctions for a variety of Riemannian manifolds.

3.1 Finding Eigenfunctions using the Rayleigh Quotient

Let us begin this section by reviewing how to find the eigenfunctions. Recall that $L^2(M)$ is the Hilbert space of square Lebesgue integrable functions:

$$L^2(M) = \left\{ \phi : M \rightarrow \mathbb{R} \mid \int_M \phi^2(x) \, d\text{vol}_x < \infty \right\} \quad (3.4)$$

with the L^2 -inner product and L^2 -norm,

$$\langle \phi, \tilde{\phi} \rangle_{L^2(M)} = \int_M \phi(x)\tilde{\phi}(x) \, d\text{vol}_x \quad \text{and} \quad \|\phi\|_{L^2}^2 = \int_M \phi^2(x) \, d\text{vol}_x. \quad (3.5)$$

To find the eigenfunctions, we begin with $\lambda_0 = 0$ and its constant eigenfunction, $\phi_0(x) = \phi_0$, which is normalized by

$$1 = \|\phi_0\|_{L^2(M)}^2 = \int_M \phi_0^2 \, d\text{vol} = \phi_0^2 \text{Vol}(M) \quad \implies \quad \phi_0(x) = \text{Vol}(M)^{-\frac{1}{2}}. \quad (3.6)$$

We find the rest of the eigenvalues iteratively using Rayleigh's Quotient:

$$\lambda_j = \inf \int_M |\nabla \phi(y)|^2 \, d\text{vol} \quad (3.7)$$

where the infimum is taken over all $\phi \in L^2(M)$ such that

$$\|\phi\|_{L^2} = 1 \quad \text{and} \quad \langle \phi, \phi_k \rangle_{L^2} = 0 \quad \text{for } k = 0, \dots, (j-1). \quad (3.8)$$

Let us review why the above Rayleigh's Quotient method works. Observe first that the conditions on ϕ in (3.8) are designed to guarantee an L^2 orthonormal collection. Next recall Green's Theorem

$$\int_M g(\nabla \phi(y), \nabla \tilde{\phi}) \, d\text{vol} = - \int_M \phi(y) \Delta \tilde{\phi} \, d\text{vol}. \quad (3.9)$$

Observe that any ϕ_j such that $\Delta \phi_j = -\lambda_j \phi_j$ with $\lambda_j > 0$ has $\langle \phi_j, \phi_k \rangle_{L^2} = 0$ for $\lambda_k < \lambda_j$ because

$$\begin{aligned} \langle \phi_j, \phi_k \rangle_{L^2} &= \int_M \phi_j(x)\phi_k(x) \, d\text{vol}_x = \frac{1}{\lambda_j} \int_M \Delta \phi_j(x)\phi_k(x) \, d\text{vol}_x \\ &= \frac{1}{\lambda_j} \int_M \phi_j(x)\Delta \phi_k(x) \, d\text{vol}_x \\ &= \frac{\lambda_k}{\lambda_j} \int_M \phi_j(x)\phi_k(x) \, d\text{vol}_x = \frac{\lambda_k}{\lambda_j} \langle \phi_j, \phi_k \rangle_{L^2}. \end{aligned}$$

If we rescale ϕ_j so that $\|\phi_j\|_{L^2} = 1$, then it achieves the infimum in (3.7):

$$\int_M |\nabla \phi_j(y)|^2 d\text{vol}_y = - \int_M \phi_j(y) \Delta \phi_j(y) d\text{vol}_y = \lambda_j \int_M \phi_j(y) \phi_j(y) d\text{vol}_y = \lambda_j.$$

It is not difficult to prove that the collection of $\phi : M \rightarrow \mathbb{R}$ which achieve the infimum in (3.7) is exactly the finite dimensional eigenspace for λ_j . We then choose an L^2 orthonormal basis of eigenfunctions for each eigenspace.

Finally we see that a complete collection of eigenfunctions found using this method forms a basis for $L^2(M)$ by proving that they span $L^2(M)$. That is, every $f \in L^2(M)$ has a Fourier Expansion:

$$f(x) = \lim_{N \rightarrow \infty} f_N(x) \text{ where } f_N(x) = \sum_{j=0}^N \langle \phi_j, f \rangle_{L^2(M)} \phi_j(x) \quad (3.10)$$

which converges in $L^2(M)$. See for example the Chavel textbook [22] for more details.

3.2 The Sturm-Liouville Decomposition of the Heat Kernel

If we wish to solve the heat equation with initial data from (3.10),

$$\partial_t u_N(x, t) = \Delta u_N(x, t) \text{ and } \lim_{t \rightarrow 0^+} u_N(x, t) = f_N(x), \quad (3.11)$$

we can solve one term at a time as in (3.3) to see that

$$\begin{aligned} u_N(x, t) &= \sum_{j=0}^N \langle \phi_j, f \rangle_{L^2} e^{-\lambda_j t} \phi_j(x) \\ &= \sum_{j=0}^N \int_M \phi_j(y) f(y) d\text{vol}_y e^{-\lambda_j t} \phi_j(x) \\ &= \int_M f(y) \sum_{j=0}^N e^{-\lambda_j t} \phi_j(x) \phi_j(y) d\text{vol}_y \\ &= \int_M f(y) h_t^N(x, y) d\text{vol}_y \end{aligned}$$

where

$$h_t^N(x, y) = \sum_{j=0}^N e^{-\lambda_j t} \phi_j(x) \phi_j(y). \quad (3.12)$$

To complete the proof of the Sturm-Liouville Decomposition, we must cautiously take $N \rightarrow \infty$, and show $u_N(x, t)$ converges to a solution $u(x, t)$ of the heat equation

$$\partial_t u(x, t) = \Delta u(x, t) \text{ and } u(x, 0) = f(x). \quad (3.13)$$

and

$$u(x, t) = \int_M f(y) (\lim_{N \rightarrow \infty} h_t^N(x, y)) dvol_y \quad (3.14)$$

so that

$$\lim_{N \rightarrow \infty} h_t^N(x, y) = h_t(x, y) \quad (3.15)$$

with convergence in $L^2(M \times M)$. See Chavel's text [22] for details.

The trace of the heat kernel is then defined naturally as follows:

$$\int_{x \in M} h_t(x, x) dvol_x = \sum_{j=0}^{\infty} e^{-\lambda_j t}. \quad (3.16)$$

3.3 Truncating the Sturm-Liouville Expansion of the Heat Kernel

Let us consider how large we must take N in order to estimate the heat kernel with the truncated sum

$$h_t^N(x, y) = \sum_{j=0}^N e^{-\lambda_j t} \phi_j(x) \phi_j(y). \quad (3.17)$$

We will define the *uniform Sturm-Liouville number*, $N_{t, \epsilon}$ to be the smallest natural number such that

$$|h_t(x, y) - h_t^N(x, y)| < \epsilon \quad \forall N \geq N_{t, \epsilon}. \quad (3.18)$$

We also define the L^2 *Sturm-Liouville number*, N_{t, ϵ, L^2} to be the smallest natural number such that $\forall N \geq N_{t, \epsilon, L^2}$ we have

$$\|h_t(x, y) - h_t^N(x, y)\|_{L^2(M \times M)} = \int_{M \times M} (h_t(x, y) - h_t^N(x, y))^2 dvol_x dvol_y < \epsilon. \quad (3.19)$$

The L^2 *Sturm-Liouville number* is easy to estimate because $\|\phi_j(x)\|_{L^2(M)} = 1$ and

$$\begin{aligned} \|\phi_j(x) \phi_j(y)\|_{L^2(M \times M)}^2 &= \int_M \int_M (\phi_j(x) \phi_j(y))^2 dvol_x dvol_y \\ &= \int_M \phi_j^2(x) dvol_x \cdot \int_M \phi_j^2(y) dvol_y \\ &= \|\phi_j\|_{L^2(M)}^2 \cdot \|\phi_j\|_{L^2(M)}^2 = 1. \end{aligned}$$

So, by the L^2 triangle inequality,

$$\|h_t(x, y) - h_t^N(x, y)\|_{L^2(M \times M)} \leq \sum_{j=N+1}^{\infty} e^{-\lambda_j t} \|\phi_j(x)\phi_j(y)\|_{L^2(M \times M)} = \sum_{j=N+1}^{\infty} e^{-\lambda_j t}. \quad (3.20)$$

This last sum is the “tail of the trace of the heat kernel” so it converges to 0 as $N \rightarrow \infty$ and can even be estimated knowing only information about the eigenvalues of the manifold.

3.4 Eigenvalues and Eigenfunctions on a Standard Circle

On a circle, \mathbb{S}^1 parametrized by $\theta \in [0, 2\pi)$ with Laplacian is

$$\Delta = \partial_{\theta}^2. \quad (3.21)$$

The zeroth eigenvalue is $\lambda_0 = 0$ with $\phi_0 = (2\pi)^{-\frac{1}{2}}$ by (3.6). The first eigenvalue is $\lambda_1 = 1$ and the eigenspace for this eigenvalue consists of periodic eigenfunctions of the form $A \sin(\theta - \theta_0)$ where $\theta_0 \in [0, 2\pi)$. This eigenspace is 2 dimensional $\lambda_2 = \lambda_1 = 1$ and we say this eigenvalue has multiplicity 2. We can choose two orthonormal eigenfunctions

$$\phi_1(\theta) = A_1 \sin(\theta) \quad \text{and} \quad \phi_2(\theta) = A_1 \sin(\theta - \pi/2) = A_1 \cos(\theta) \quad (3.22)$$

where

$$1 = \int_0^{2\pi} A_1^2 \sin^2(\theta) d\theta = \left(\frac{1}{2}\right) \int_0^{2\pi} A_1^2 d\theta = \left(\frac{1}{2}\right) 2\pi A_1^2, \quad \text{so} \quad A_1 = \pi^{-\frac{1}{2}}. \quad (3.23)$$

Continuing forward we find that $\lambda_{2j} = \lambda_{2j-1} = j^2$ with

$$\phi_{2j-1}(\theta) = A_j \sin(j\theta) \quad \text{and} \quad \phi_{2j}(\theta) = A_j \cos(j\theta) \quad (3.24)$$

where

$$1 = \int_0^{2\pi} A_j^2 \sin^2(j\theta) d\theta = \left(\frac{1}{2}\right) \int_0^{2\pi} A_j^2 d\theta = \pi A_j^2 \quad \text{so} \quad A_j = \pi^{-\frac{1}{2}}. \quad (3.25)$$

So following (3.1), the truncated Sturm-Liouville expansion for the heat kernel \mathbb{S}^1 is

$$h_t^{2N+1}(\theta_1, \theta_2) = \frac{1}{2\pi} + \frac{1}{\pi} \sum_{j=1}^N e^{-tj^2} (\sin(j\theta_1) \sin(j\theta_2) + \cos(j\theta_1) \cos(j\theta_2)). \quad (3.26)$$

The trace of the heat kernel on \mathbb{S}^1 is

$$\int_{\theta=0}^{2\pi} h_t(\theta, \theta) d\theta = 1 + 2 \sum_{j=1}^{\infty} e^{-j^2 t}. \quad (3.27)$$

Since the eigenfunctions of \mathbb{S}^1 are uniformly bounded, $|\phi_j(x)| \leq \pi^{-1/2}$, the Sturm-Liouville expansion converges as quickly as the trace of heat kernel and we can use the same estimates to compute the uniform and the L^2 Sturm-Liouville numbers.

If we try to apply Varadhan's formula to the truncated Sturm-Liouville expansion for the heat kernel, we see that as $t \rightarrow 0$, the truncated heat kernels, $h_t^{2N+1}(\theta_1, \theta_2)$, converge to

$$(2\pi)^{-1} + (\pi)^{-1} \sum_{j=1}^N (\sin(j\theta_1) \sin(j\theta_2) + \cos(j\theta_1) \cos(j\theta_2)). \quad (3.28)$$

So we end up with nothing:

$$\lim_{t \rightarrow 0} -4t \log h_t^{2N+1}(\theta_1, \theta_2) = 0. \quad (3.29)$$

In order to estimate Varadhan's formula using the Sturm-Liouville expansion with an error less than $\varepsilon > 0$, we would have to first fix t_ε small enough that

$$|(-4t \log h_t(\theta_1, \theta_2)) - d_{\mathbb{S}^1}^2(\theta_1, \theta_2)| < \varepsilon/2 \quad (3.30)$$

and then for fixed $t = t_\varepsilon$ take N_t large enough that

$$|(-4t \log h_t^{N_t}(\theta_1, \theta_2)) - (-4t \log h_t(\theta_1, \theta_2))| < \varepsilon/2. \quad (3.31)$$

This would then give us

$$|d_{\mathbb{S}^1}^2(\theta_1, \theta_2) + 4t \log h_t^{N_t}(\theta_1, \theta_2)| < \varepsilon \quad (3.32)$$

for $\theta_1, \theta_2 \in \mathbb{S}^1$, $t = t_{\varepsilon/2}$, and $N = N_{t_\varepsilon, \varepsilon/2}$. To get a better estimate, we need to take t smaller first and then N larger.

3.5 Eigenvalues and Eigenfunctions on a Rescaled Circle

Let us consider a circle, \mathbb{S}_R^1 , where $x = R\theta \in [0, 2\pi R)$ with $g_R = dx^2 = R^2 d\theta^2$. The Laplacian rescales

$$\Delta_R = \partial_x^2 = R^{-2} \partial_\theta^2 \quad (3.33)$$

and the volume is $2\pi R$. We find that $\lambda_0 = 0$ with $\phi_0 = (2\pi R)^{-1/2}$ by (3.6). The rest of the eigenvalues come with multiplicity two:

$$\lambda_{2j} = \lambda_{2j-1} = j^2/R^2 \quad (3.34)$$

with L^2 orthonormal eigenfunctions

$$\begin{aligned} \phi_{2j-1}(\theta) &= A_j \sin(j\theta) = A_j \sin(jx/R) \\ \phi_{2j}(\theta) &= A_j \cos(j\theta) = A_j \cos(jx/R) \end{aligned}$$

where

$$1 = \int_0^{2\pi} A_j^2 \sin^2(j\theta) R d\theta = \frac{1}{2} \int_0^{2\pi} A_j^2 R d\theta = 2\pi A_j^2 R/2 \text{ so } A_j = (\pi R)^{-1/2}. \quad (3.35)$$

So the Sturm-Liouville expansion for \mathbb{S}_R^1 is

$$h_t^R(\theta_1, \theta_2) = \frac{1}{2\pi R} + \frac{1}{\pi R} \sum_{j=1}^{\infty} e^{-tj^2/R^2} (\sin(j\theta_1) \sin(j\theta_2) + \cos(j\theta_1) \cos(j\theta_2)). \quad (3.36)$$

Note that this rescales like (2.13). The trace of the heat kernel is then

$$\int_{\theta=0}^{2\pi} h_t^R(\theta, \theta) R d\theta = 1 + 2 \sum_{j=1}^{\infty} e^{-tj^2/R^2}. \quad (3.37)$$

Observe that as $R \rightarrow 0$ this converges to 0 as expected.

3.6 Eigenvalues and Eigenfunctions on Tori

Given any $A \leq B$, the flat torus formed by taking the isometric product of rescaled circles,

$$\mathbb{S}_A^1 \times \mathbb{S}_B^1 = (\mathbb{S}^1 \times \mathbb{S}^1, dx^2 + dy^2), \quad (3.38)$$

where $x = A\theta$ and $y = B\varphi$ has the following Laplacian,

$$\Delta_{A,B} = \partial_x^2 + \partial_y^2 = A^{-2} \partial_\theta^2 + B^{-2} \partial_\varphi^2. \quad (3.39)$$

Its volume is $4\pi^2 AB$ so, by (3.6), for $\lambda_0 = 0$ the eigenfunction is $\phi_0 = (4\pi^2 AB)^{-1/2}$. The rest of the eigenfunctions come in sets of four as follows:

$$\{\cos(j\theta) \cos(k\varphi), \cos(j\theta) \sin(k\varphi), \sin(j\theta) \cos(k\varphi), \sin(j\theta) \sin(k\varphi)\} \quad (3.40)$$

which we can rescale by constants $A_{j,k}$ to be L^2 orthonormal. Each set has the same eigenvalue

$$\lambda_{j,k} = j^2/A^2 + k^2/B^2. \quad (3.41)$$

The ordering of these eigenvalues depends on the values of A and B , so first lets examine the values of $A_{j,k}$.

When $j = 0$ and $k = 0$, we recover $\lambda_{0,0} = \lambda_0 = 0$ and see by (3.6) that there is a single L^2 -orthonormal eigenfunction,

$$\lambda_{0,0} = 0 \text{ with } V_{0,0} = \{A_{0,0} \cos(0) \cos(0)\} \quad (3.42)$$

where $A_{0,0} = \frac{1}{\sqrt{(2\pi A)(2\pi B)}} = \frac{1}{2\pi \sqrt{AB}}$.

When $j > 0$ and $k = 0$, we also get $\lambda_{j,0} = j^2/A^2$ with two L^2 -orthonormal eigenfunctions,

$$\lambda_{j,0} = \frac{j^2}{A^2} \text{ with } V_{j,0} = \{A_{j,0} \cos(j\theta) \cos(0), A_{j,0} \sin(j\theta) \cos(0)\} \quad (3.43)$$

where $A_{j,0} = \frac{1}{\pi \sqrt{2AB}}$ because

$$1 = \int_0^{2\pi} \int_0^{2\pi} A_{j,0}^2 \sin^2(j\theta) AB \, d\theta d\varphi = A_{j,0}^2 (\frac{1}{2}) (2\pi A)(2\pi B). \quad (3.44)$$

When $j = 0$ and $k > 0$, we get $\lambda_{0,k} = k^2/B^2$ with two L^2 -orthonormal eigenfunctions,

$$\lambda_{0,k} = \frac{k^2}{B^2} \text{ with } V_{0,k} = \{A_{0,k} \cos(0) \cos(k\varphi), A_{0,k} \cos(0) \sin(k\varphi)\} \quad (3.45)$$

where $A_{0,k} = \frac{1}{\pi \sqrt{2AB}}$.

When both $j, k > 0$, we get $\lambda_{j,k} = j^2/A^2 + k^2/B^2$ with four L^2 -orthonormal eigenfunctions as in (3.40),

$$\lambda_{j,k} = \frac{j^2}{A^2} + \frac{k^2}{B^2} \text{ with } V_{j,k} = \{A_{j,k} \cos(j\theta) \cos(k\varphi), \dots, A_{j,k} \sin(j\theta) \sin(k\varphi)\} \quad (3.46)$$

where $A_{j,k} = \frac{1}{\pi \sqrt{AB}}$, because

$$\begin{aligned} 1 &= \int_0^{2\pi} \int_0^{2\pi} A_{j,k}^2 \sin^2(j\theta) \sin^2(k\varphi) AB \, d\theta d\varphi \\ &= A_{j,k}^2 \cdot \int_0^{2\pi} \sin^2(j\theta) A \, d\theta \cdot \int_0^{2\pi} \sin^2(k\varphi) B \, d\varphi \\ &= A_{j,k}^2 \cdot (2\pi A(\frac{1}{2})) \cdot (2\pi B(\frac{1}{2})) = A_{j,k}^2 \pi^2 AB. \end{aligned}$$

When $A = B = 1$, the eigenvalues are ordered as follows with the eigenspaces:

$$\begin{aligned}
 \lambda_{0,0} &= 0^2 + 0^2 = 0 & V_{0,0} \text{ of multiplicity } 1 \\
 \lambda_{1,0} &= \lambda_{0,1} = 1^2 + 0^2 = 1 & V_{1,0} \cup V_{0,1} \text{ of multiplicity } 2 + 2 = 4 \\
 \lambda_{1,1} &= 1^2 + 1^2 = 2 & V_{1,1} \text{ of multiplicity } 4 \\
 \lambda_{2,0} &= \lambda_{0,2} = 2^2 + 0^2 = 4 & V_{2,0} \cup V_{0,2} \text{ of multiplicity } 2 + 2 = 4 \\
 \lambda_{2,1} &= \lambda_{1,2} = 2^2 + 1^2 = 5 & V_{2,1} \cup V_{1,2} \text{ of multiplicity } 4 + 4 = 8 \\
 \lambda_{2,2} &= 2^2 + 2^2 = 8 & V_{2,2} \text{ of multiplicity } 4 \\
 \lambda_{3,0} &= \lambda_{0,3} = 3^2 + 0^2 = 9 & V_{3,0} \cup V_{0,3} \text{ of multiplicity } 2 + 2 = 4.
 \end{aligned}$$

When $A = 1$ and $B = \frac{1}{2}$, the eigenvalues $\lambda_{j,k} = j^2 + 4k^2$ are ordered quite differently with different multiplicities:

$$\begin{aligned}
 \lambda_{0,0} &= 0^2 + 0^2 = 0 & V_{0,0} \text{ of multiplicity } 1 \\
 \lambda_{1,0} &= 1^2 + 4 \cdot 0^2 = 1 & V_{1,0} \text{ of multiplicity } 2 \\
 \lambda_{0,1} &= \lambda_{2,0} = 0^2 + 4 \cdot 1^2 = 4 & V_{0,1} \cup V_{2,0} \text{ of multiplicity } 2 + 2 = 4 \\
 \lambda_{1,1} &= 1^2 + 4 \cdot 1^2 = 5 & V_{1,1} \text{ of multiplicity } 4 \\
 \lambda_{2,1} &= 2^2 + 4 \cdot 1^2 = 8 & V_{2,1} \text{ of multiplicity } 4 \\
 \lambda_{3,0} &= 3^2 + 4 \cdot 0^2 = 9 & V_{3,0} \text{ of multiplicity } 2.
 \end{aligned}$$

This ordering of the eigenfunctions becomes important when we try to apply these eigenfunctions to Manifold Learning.

3.7 Eigenvalues and Eigenfunctions on a Thin Torus

Let us consider a thin torus, $\mathbb{S}_A^1 \times \mathbb{S}_B^1$ with $A = 1$ and $B = 1/10$. Here the eigenvalues are $\lambda_{j,k} = j^2 + 100k^2$. Following our method from Section 3.6 we have the following first 23 eigenvalues (counting multiplicity) and their eigenspaces:

$$\begin{aligned}
 \lambda_{0,0} &= 0^2 + 0^2 = 0 & V_{0,0} \text{ of multiplicity } 1 \\
 \lambda_{j,0} &= j^2 + 100 \cdot 0^2 = j^2 & V_{j,0} \text{ of multiplicity } 2 \text{ for } j = 1, \dots, 9 \\
 \lambda_{0,1} &= \lambda_{10,0} = 100 & V_{0,1} \cup V_{10,0} \text{ of multiplicity } 2 + 2 = 4.
 \end{aligned}$$

Note that the eigenspaces $V_{j,0}$ as in (3.43) do not have functions depending on φ . Thus the truncated Sturm-Liouville expansion with $N = 19$ for the heat kernel of this thin torus does not depend on φ :

$$h_t^{19}((\theta_1, \varphi_1), (\theta_2, \varphi_2)) = \frac{1}{4\pi^2 B} + \frac{1}{2\pi^2 B} \sum_{j=1}^9 e^{-j^2 t} (\sin(j\theta_1) \sin(j\theta_2) + \cos(j\theta_1) \cos(j\theta_2))$$

because $(1/(2\pi\sqrt{AB}))^2 = 1/(4\pi B^2)$ when $A = 1$.

Note that this $N = 19$ truncated Sturm-Liouville expansion for the heat kernel of a thin torus is equal to the $N = 19$ truncated Sturm-Liouville expansion for the heat kernel of a circle, \mathbb{S}^1 , that we found in (3.26) rescaled by $1/(2\pi B)$. *Intuitively the truncated heat kernel views this thin torus as almost the same space as a circle.* This becomes important when using eigenfunctions for Manifold Learning: sometimes it can be advantageous to view a thin torus as almost the same as a circle and other times it is better to distinguish them.

3.8 Eigenvalues and Eigenfunctions on a Sphere

The eigenfunctions of the sphere, \mathbb{S}^2 are called Laplace's spherical harmonics. They are explained beautifully in many sources including wikipedia: [Spherical Harmonics](#). The eigenvalues take the form can be written as

$$\ell(\ell + 1) \text{ with eigenfunctions } Y_\ell^m = \cos(m\varphi)P_\ell^m(\cos(\theta)) \quad (3.47)$$

where $m = -\ell, -\ell + 1, \dots, 0, \dots, \ell - 1, \ell$ and P_ℓ^m is a Legendre Polynomial where here we are using the coordinates $\varphi \in [0, \pi]$ and $\theta \in [0, 2\pi]$ so that

$$g_{\mathbb{S}^2} = d\varphi^2 + \sin^2(\varphi)d\theta^2. \quad (3.48)$$

The first eigenvalue $\lambda_0 = 0(1) = 0$ with $\phi_0 = (4\pi)^{-1/2}$ by (3.6). The next eigenvalue is $\lambda_1 = 1(2) = 2$ with eigenfunction

$$\phi_1(\theta, \varphi) = A_1 \cos(\varphi) = A_1 Z(\theta, \varphi) \quad (3.49)$$

where $Z(\theta, \varphi)$ is the Z coordinate of the standard Riemannian isometric embedding of the sphere into Euclidean space,

$$X(\theta, \varphi) = \sin(\varphi) \cos(\theta) \quad Y(\theta, \varphi) = \sin(\varphi) \sin(\theta) \quad Z(\theta, \varphi) = \cos(\varphi). \quad (3.50)$$

By symmetry we can see that $X(\theta, \varphi)$ and $Y(\theta, \varphi)$ also give eigenfunctions, so

$$\phi_2(\theta, \varphi) = A_1 X(\theta, \varphi) \quad \text{and} \quad \phi_3(\theta, \varphi) = A_1 Y(\theta, \varphi) \quad (3.51)$$

with $\lambda_2 = \lambda_3 = \lambda_1 = 2$. To compute A_1 , we can use the symmetry and the fact that $X^2 + Y^2 + Z^2 = 1$ to see by (3.49) and (3.51) that

$$\begin{aligned} 1 &= \int_{\mathbb{S}^2} \phi_1^2 d\text{vol} = \frac{1}{3} \int_{\mathbb{S}^2} (\phi_1^2 + \phi_2^2 + \phi_3^2) d\text{vol} \\ &= \frac{1}{3} \int_{\mathbb{S}^2} A_1^2 (X^2 + Y^2 + Z^2) d\text{vol} = \frac{1}{3} \int_{\mathbb{S}^2} A_1^2 (1) d\text{vol} = \frac{4\pi A_1^2}{3}. \end{aligned}$$

Thus $A_1 = \sqrt{3/(4\pi)}$. It is an exercise to check that these are orthonormal.

The $N = 4$ truncated Sturm-Liouville expansion of the heat kernel for \mathbb{S}^2 can be viewed as a dot product of the embedding map (3.50):

$$\begin{aligned} h_t^4(x_1, x_2) &= \frac{1}{4\pi} + \frac{3}{4\pi} e^{-2t} (X(x_1)X(x_2) + Y(x_1)Y(x_2) + Z(x_1)Z(x_2)) \\ &= \frac{1}{4\pi} + \frac{3}{4\pi} e^{-2t} \cos(d_{\mathbb{S}^2}(x_1, x_2)). \end{aligned}$$

This dependence on the distance is a consequence of the symmetry of \mathbb{S}^2 and is unrelated Varadhan's formula because it holds without taking $N \rightarrow \infty$ and $t \rightarrow 0$.

3.9 Estimates on the First Eigenvalue

In 1969, Cheeger found the following lower bound for the first nonzero eigenvalue of the Laplacian:

Theorem 3.1. Let M be a compact manifold with no boundary. Then,

$$\lambda_1 \geq \frac{1}{4} h^2(M). \quad (3.52)$$

Here $h(M)$ is Cheeger's Constant:

$$h(M) = \inf \left\{ \frac{\text{Vol}_{n-1}(\partial U)}{\min\{\text{Vol}_n(U), \text{Vol}_n(M \setminus U)\}} \right\} \quad (3.53)$$

where the infimum is taken over all domains $U \subset M$.

This was proven using Rayleigh's Quotient and coarea formulas in a clever way. See, for example, the proof in [83].

In 1975 S-T Yau extended Cheeger's result and applied it to prove a lower bound on this eigenvalue in terms of the Ricci curvature on the manifold in [96].

Theorem 3.2. When $\text{Ricci} \geq (n-1)H$, then

$$\lambda_1 \geq \frac{1}{4} I^2(M) \quad (3.54)$$

where

$$\begin{aligned} I(M)^{-1} &\leq \omega_{n-1} \frac{\text{Diam}(M)}{\text{Vol}(M)} \int_0^{\text{Diam}(M)} \left(|H|^{-\frac{1}{2}} \sin(|H|^{\frac{1}{2}} r) \right)^{n-1} dr \quad \text{when } H > 0, \\ I(M)^{-1} &\leq \omega_{n-1} \frac{\text{Diam}(M)}{\text{Vol}(M)} \int_0^{\text{Diam}(M)} (r)^{n-1} dr \quad \text{when } H = 0, \\ I(M)^{-1} &\leq \omega_{n-1} \frac{\text{Diam}(M)}{\text{Vol}(M)} \int_0^{\text{Diam}(M)} \left(|H|^{-\frac{1}{2}} \sinh(|H|^{\frac{1}{2}} r) \right)^{n-1} dr \quad \text{when } H < 0. \end{aligned}$$

In 1980 Buser proved Cheeger's inequality is sharp by showing that any manifold has a Riemannian metric tensor which achieves this inequality [20]. In 1982, Buser [21] proved the following uniform estimate:

Theorem 3.3. If M^n has $\text{Ricci} \geq (n - 1)H$ where $H < 0$ then

$$\lambda_1(M) \leq c(n) \left(h(M) |H|^{1/2} + h^2(M) \right). \quad (3.55)$$

Buser presented two proofs in [21], the second of which has a Jacobi field argument which is especially nice to read for those who enjoy do Carmo's textbook [46]. Buser also studied a variety of examples in these papers as well.

3.10 Estimating Eigenvalues on Families of Manifolds

Explicit computations of the eigenfunctions and eigenvalues for many Riemannian manifolds can be found in the French text of Berger, Gauduchon, and Mazet [17]. See also Chavel for many examples with constant negative curvature [22].

S.Y. Cheng gave upper estimates on all the eigenvalues of a Riemannian manifold with $\text{Ricci} \geq (n - 1)H$ [28]. In 1980, Li and Yau improved this result, proving many results including

$$\lambda_k \leq C_{H,D,n} \left(\frac{H+1}{V} \right)^{2/n} + C'_{H,n} \quad (3.56)$$

for any n dimensional Riemannian manifold M^n with $\text{Ricci} \geq (n - 1)H$, $\text{Diam}(M) \leq D$ and $\text{Vol}(M) \geq V$ [68].

There are also lower bounds found by Cheng-Li in [29] depending on the Sobolev constant. See also the work of Debiard in [42] and of Hess-Schrader-Uhlenbrock in [57]. For more details see the text of Schoen and Yau [83].

4 Classes of Riemannian Manifolds and Convergence

In Sections 2.5 and 3.10 we have seen that entire classes of families of Riemannian manifolds have uniform bounds on their heat kernels and eigenvalues. This is closely related to the various notions of convergence of Riemannian manifolds, their corresponding compactness theorems, and how the heat kernels and eigenvalues behave on converging sequences. To be more clear, suppose we have a notion of convergence of Riemannian manifolds, $M_j \rightarrow M_\infty$, then we say

a class, \mathcal{M} , of Riemannian manifolds is *compact* with respect to this notion of convergence if the following holds:

$$M_j \subset \mathcal{M} \implies \exists \text{ a subseq } M_{j_k} \text{ s.t. } M_{j_k} \rightarrow M_\infty \in \mathcal{M}. \quad (4.1)$$

The class \mathcal{M} is only *precompact* if the subsequence converges to M_∞ that is not in \mathcal{M} .

For many notions of convergence, $M_j \rightarrow M_\infty$, given any pair of points in the limit space, $p, q \in M_\infty$, there are $p_j, q_j \in M_j$ that have $p_j \rightarrow p_\infty$ and $q_j \rightarrow q_\infty$. We say a notion of convergence *preserves distances* if $M_j \rightarrow M_\infty$ implies

$$d_j(p_j, q_j) \rightarrow d_\infty(p, q). \quad (4.2)$$

We say a notion of convergence *preserves the heat kernel* if $M_j \rightarrow M_\infty$ implies

$$\forall t > 0, h_t^{M_j}(p_j, q_j) \rightarrow h_t^{M_\infty}(p, q). \quad (4.3)$$

We say there is *continuity of the eigenvalues* if $M_j \rightarrow M_\infty$ implies

$$\forall k \in \mathbb{N} \lambda_k(M_j) \rightarrow \lambda_k(M_\infty). \quad (4.4)$$

Sometimes we only have semicontinuity of the eigenvalues.

If we have a family of manifolds, \mathcal{M} , and a notion of convergence which makes this family compact and this notion of convergence implies continuity of the eigenvalues, then we have uniform bounds on the eigenvalues. Otherwise, we could take a sequence of $M_j \in \mathcal{M}$ whose eigenvalues are diverging, find the converging subsequence and use continuity of the eigenvalues on the subsequence to reach a contradiction. Similarly, if the notion of convergence preserves the heat kernel, then it has uniform bounds on the heat kernel.

Below we will discuss various notions of convergence and their relationship with heat kernels and eigenvalues.

4.1 Smooth, C^0 , $C^{k,\alpha}$, and C^k Convergence

Recall that a function $h : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ is in $C^0(U)$ if it is continuous and the C^0 norm is

$$|h|_{C^0} = \sup_{x \in U} |h(x)|. \quad (4.5)$$

We say that $h \in C^{0,\alpha}(U)$ if it's Hölder norm is bounded:

$$|h|_{C^{0,\alpha}(U)} = \sup_{x \neq y \in U} \frac{|h(x) - h(y)|}{|x - y|^\alpha} < \infty. \quad (4.6)$$

We say that $h \in C^k(U)$ if

$$|h|_{C^k(U)} = \sup_{x \in U} \sum |\partial_{x_1}^{\beta_1} \cdots \partial_{x_n}^{\beta_n} h(x)| < \infty, \quad (4.7)$$

where the sum is over $(\beta_1, \dots, \beta_n)$ such that $\beta_1 + \cdots + \beta_n \leq k$. We say that $h \in C^{k,\alpha}(U)$ if

$$|h|_{C^{k,\alpha}(U)} = |h|_{C^k(U)} + \sum |\partial_{x_1}^{\beta_1} \cdots \partial_{x_n}^{\beta_n} h(x)|_{C^{0,\alpha}(U)} < \infty, \quad (4.8)$$

where the sum is over $(\beta_1, \dots, \beta_n)$ such that $\beta_1 + \cdots + \beta_n = k$.

We say an atlas, $\{\varphi_i\}$, of charts on M is C^k or $C^{k,\alpha}$ if the transition maps are C^k or $C^{k,\alpha}$ respectively. If we fix a C^k or $C^{k,\alpha}$ atlas, $\{\varphi_i\}$, of charts on M , then with respect to that atlas we can define:

$$|h|_{C^k(M),\{\varphi\}} = \sup |h \circ \varphi|_{C^k(U)} \quad \text{or} \quad |h|_{C^{k,\alpha}(M),\{\varphi\}} = \sup |h \circ \varphi|_{C^{k,\alpha}(U)} \quad (4.9)$$

where the sup is over all coordinate charts $\varphi : U \rightarrow M$ in the atlas.

A function between compact manifolds, $H : M_1 \rightarrow M_2$ is C^k , if there exists a C^k atlas of charts on M_1 and on M_2 such that the function composed with these charts is C^k respectively. It is a C^k diffeomorphism if it is invertible and both it and its inverse are C^k . It is a $C^{k,\alpha}$ diffeomorphism if everything here is $C^{k,\alpha}$. It is a smooth or C^∞ diffeomorphism if it is C^k for all $k \in \mathbb{N}$.

We say $(M_j, g_j) \rightarrow (M_\infty, g_\infty)$ in the C^k sense if $H_j : M_\infty \rightarrow M_j$ is a C^{k+1} diffeomorphism (which means it has $(k+1)$ continuous derivatives) and if there exists a C^k atlas on M_∞ such that for all g_∞ -unit length $u, v \in TM_\infty$ we have

$$|H_j^* g_j(u, v) - g_\infty(u, v)|_{C^k(M),\{\varphi_i\}} < \epsilon_j \rightarrow 0. \quad (4.10)$$

We say $(M_j, g_j) \rightarrow (M_\infty, g_\infty)$ in the $C^{k,\alpha}$ sense if $H_j : M_\infty \rightarrow M_j$ is a $C^{k+1,\alpha}$ diffeomorphism and if there exists a $C^{k,\alpha}$ atlas on M_∞ such that for all g_∞ -unit length $u, v \in TM_\infty$ we have

$$|H_j^* g_j(u, v) - g_\infty(u, v)|_{C^{k,\alpha}(M),\{\varphi_i\}} < \epsilon_j \rightarrow 0. \quad (4.11)$$

It converges smoothly in the C^∞ sense if all the diffeomorphisms are C^∞ and if there exists a C^∞ atlas on M_∞ such that for all g_∞ -unit length $u, v \in TM_\infty$ we have

$$|H_j^* g_j(u, v) - g_\infty(u, v)|_{C^\infty(M),\{\varphi_i\}} < \epsilon_j \rightarrow 0. \quad (4.12)$$

4.2 Smooth Convergence of Tori and Warped Tori

Note that if we consider tori $\mathbb{S}_{A_j}^1 \times \mathbb{S}_{B_j}^1$ and $A_j \rightarrow A_\infty > 0$ and $B_j \rightarrow B_\infty > 0$ then we have C^∞ convergence to $\mathbb{S}_{A_\infty}^1 \times \mathbb{S}_{B_\infty}^1$. This can be seen by taking $\theta, \varphi \in [0, 2\pi)$ and writing the metric tensors,

$$g_{\mathbb{S}_A^1 \times \mathbb{S}_B^1} = A^2 d\theta^2 + B^2 d\varphi^2, \quad (4.13)$$

and taking $H_j : \mathbb{S}_{A_\infty}^1 \times \mathbb{S}_{B_\infty}^1 \rightarrow \mathbb{S}_{A_j}^1 \times \mathbb{S}_{B_j}^1$ preserving θ and φ . Then

$$|H_j^* g_j(u, v) - g_\infty(u, v)|_{C^0} = |(A_j^2 - A_\infty^2)u(\theta)v(\theta) + (B_j^2 - B_\infty^2)u(\varphi)v(\varphi)|_{C^0} \rightarrow 0 \quad (4.14)$$

It is easy to see there is C^k convergence as well.

A sequence of warped product metric tensors on a torus converges in the C^k sense

$$g_j = d\varphi^2 + f_j(\varphi)^2 d\theta^2 \rightarrow g_\infty = d\varphi^2 + f_\infty(\varphi)^2 d\theta^2 \quad (4.15)$$

if and only if $|f_j - f_\infty|_{C^k} \rightarrow 0$. On tori we must assume the warping factors are periodic, $f_j(\varphi) = f_j(\varphi + 2\pi)$, and positive, $f_j > 0$ and $f_\infty > 0$, to have a diffeomorphism. If we only have $f_j(0) = f_j(\pi)$ without requiring the derivatives to match as implied by periodicity, then we only have a C^0 metric tensor and C^0 convergence. To have a C^k metric tensor, we need k derivatives to match:

$$f_j(0) = f_j(\pi) \quad f_j'(0) = f_j'(\pi) \quad f_j''(0) = f_j''(\pi) \quad \dots \quad f_j^{(k)}(0) = f_j^{(k)}(\pi). \quad (4.16)$$

To have C^k convergence, we need all these derivatives to converge

$$f_j^{(i)}(0) \rightarrow f_\infty^{(i)}(0) \text{ for } i = 0, \dots, k.$$

A sequence of warped product metric tensors on a sphere will have (4.15) with $f_j(0) = f_j(\pi) = 0$ and $f_\infty(0) = f_\infty(\pi)$. To have a C^k metric tensor across the poles, we need k derivatives to match that of the standard sphere:

$$\begin{aligned} f_j(0) &= f_j(\pi) = \sin(0) = 0 \\ f_j'(0) &= f_j'(\pi) = \sin'(0) = 1 \\ f_j''(0) &= f_j''(\pi) = \sin''(0) = 0 \\ &\dots \\ f_j^{(k)}(0) &= f_j^{(k)}(\pi) = \sin^{(k)}(0). \end{aligned}$$

To have C^k convergence, these derivatives must converge

$$f_j^{(i)}(0) \rightarrow f_\infty^{(i)}(0) \text{ for } i = 0, \dots, k.$$

4.3 Convergence of Distances between Points

Distances are preserved under C^0 convergence of $M_j \rightarrow M_\infty$. To see this, first we show that for any curve $C : [0, 1] \rightarrow M$ such that $C(0) = p_\infty$ and $C(1) = q_\infty$ we have $C_j = C \circ H_j^{-1} : [0, 1] \rightarrow M_j$ with $C_j(0) = p_j$ and $C_j(1) = q_j$. Since

$$L_{g_j}(C_j) = \int_0^1 g_j(C'_j(s), C'_j(s))^{1/2} ds = \int_0^1 (H_j^* g_j)(C'(s), C'(s))^{1/2} ds \rightarrow L_{g_\infty}(C).$$

we see that $|L_{g_j}(C_j) - L_{g_\infty}(C)| < \epsilon_j^{1/2}$. Since the distance d_j is the infimum of the length, L_j , of all curves from p_j to q_j , we have

$$d_j(p_\infty, q_\infty) \leq L_{g_j}(C_j) \leq L_{g_\infty}(C) + \epsilon_j^{1/2}. \quad (4.17)$$

Since the distance, d_∞ , is the infimum of the length, L_∞ , over all curves, C , as above

$$d_j(p_\infty, q_\infty) \leq d_\infty(p_\infty, q_\infty) + \epsilon_j^{1/2} \quad (4.18)$$

Applying this entire argument in reverse we get the desired convergence:

$$|d_\infty(p_\infty, q_\infty) - d_j(p_j, q_j)| \leq \epsilon_j^{1/2} \rightarrow 0. \quad (4.19)$$

4.4 Lipschitz Convergence

Lipschitz convergence of Riemannian manifolds is defined by viewing a Riemannian manifold, (M, g) , as a metric space (M, d_g) where d_g is the Riemannian distance. We say two Riemannian Manifolds are biLipschitz if there exists a bijection $H : M_1 \rightarrow M_2$ such that

$$\begin{aligned} \text{Lip}(H) &= \sup_{x \neq y} d_2(H(x), H(y))/d_1(x, y) < \infty \\ \text{Lip}(H^{-1}) &= \sup_{x \neq y} d_1(H^{-1}(x), H^{-1}(y))/d_2(x, y) < \infty. \end{aligned}$$

The Lipschitz distance between them is:

$$d_{Lip}((M_1, d_1), (M_2, d_2)) = \log \min\{\text{Lip}(H), \text{Lip}(H^{-1}) : H : M_1 \rightarrow M_2\}. \quad (4.20)$$

It is easy to verify that $M_j \rightarrow M_\infty$ in the C^0 sense implies $d_{Lip}(M_j, M_\infty) \rightarrow 0$.

Consider the class of smooth Riemannian manifolds which are biLipschitz to a standard sphere, \mathbb{S}^2 with Lipschitz bounds in both directions ≤ 2 . By the

Arzela-Ascoli Theorem, this class is precompact with respect to Lipschitz convergence. However, the limit spaces need not be a smooth Riemannian manifolds.

In fact M_∞ could have a conical singularity. This can be seen by considering the following sequence of functions:

$$g_j = d\varphi^2 + f_j(\varphi)^2 d\theta^2 \quad (4.21)$$

where f_j are smooth functions such that

$$f_j(\varphi) = \sin(\varphi) \text{ on } [0, \delta_j] \cup [\pi - \delta_j, \pi]$$

$$f_j(\varphi) = \sin(\varphi)/2 \text{ on } [0, 5\delta_j] \cup [\pi - 5\delta_j, \pi]$$

and monotone on $[\delta_j, 5\delta_j] \cup [\pi - 5\delta_j, \pi - \delta_j]$. Taking $\delta_j \rightarrow 0$, the Lipschitz limit is an American-football-shaped C^0 Riemannian manifold with conical singularities at the poles where $\varphi = 0$ and $\varphi = \pi$.

This sequence has Ricci and sectional curvature bounded from below by 0 but not from above and has no lower bound on the injectivity radius. However we can still apply Cheeger-Yau's theorem in (2.30) to say that the heat kernels of these manifolds are bounded from below by the heat kernel on Euclidean space. On balls which avoid the poles, the heat kernels are bounded on both sides. In fact as we shall soon see, the heat kernels converge to a function on the limit space which has many of the properties of a heat kernel.

4.5 C^2 convergence preserves Δ and $h_t(x, y)$

If $(M_j, g_j) \rightarrow (M_\infty, g_\infty)$ in the C^2 sense, then it is easy to see that their curvatures and Laplacians are preserved,

$$|\Delta^{M_j}(h \circ H_j^{-1}) - \Delta^{M_\infty} h|_{C^0} \rightarrow 0 \text{ for any } C^2 \text{ function } h : M_\infty \rightarrow \mathbb{R}. \quad (4.22)$$

because Laplacians depend only on first and second derivatives of the metric tensor.

Since solutions to the heat equation depend smoothly on the Laplace operator, the heat kernel on M_j can be shown to converge to the heat kernel on M_∞ ,

$$|h_t^{M_j}(H_j^{-1}(x), H_j^{-1}(y)) - h_t^{M_\infty}(x, y)|_{C^2} \rightarrow 0, \quad (4.23)$$

and similarly the eigenfunctions can be shown to converge

$$|\phi_k^{M_j}(H_j^{-1}(x)) - \phi_k^{M_\infty}(x)|_{C^2} \rightarrow 0 \text{ and } |\lambda_k(M_j) - \lambda_k(M_\infty)| \rightarrow 0 \quad (4.24)$$

where as usual the λ_k are counted with multiplicity.

What is quite surprising is that heat kernels and eigenvalues can converge even with far weaker notions of convergence.

4.6 Diffeomorphic Compactness Theorems

With the above definitions of convergence of manifolds we require that the sequence of manifolds be diffeomorphic in order to compare their metric tensors. So if we are studying a class or family of Riemannian manifolds, \mathcal{M} , and wish to prove a compactness theorem stating that given a sequence $M_j \in \mathcal{M}$ a subsequence converges to $M_\infty \in \mathcal{M}$, we first need to show a subsequence is pairwise diffeomorphic to one another. This can be achieved using the pigeonhole principle if we know there are only finitely many diffeomorphism types in the class \mathcal{M} .

In his 1967 doctoral dissertation [23, 24], Cheeger observed that there were only finitely many diffeomorphism types in

$$\mathcal{M}_{K_1, V}^{K_2, D} = \{M^n \mid K_1 \leq \text{sect} \leq K_2, \text{Vol}(M) \geq V, \text{Diam}(M) \leq D\} \quad (4.25)$$

by proving a uniform lower bound on injectivity radius based on K_1 , K_2 , V , and D . He also found other classes of Riemannian manifolds with similar controls.

Note that the tori

$$\mathbb{S}_A^1 \times \mathbb{S}_B^1 \in \mathcal{M}_{K_1, V}^{K_2, D} \text{ for any } K_1 \leq 0 \leq K_2 \quad (4.26)$$

if and only if

$$4\pi^2 AB \geq V \text{ and } (\pi A)^2 + (\pi B)^2 \leq D^2 \quad (4.27)$$

because they are flat with sectional curvature = 0. Notice that (4.27) implies $A, B \leq D/\pi$ which implies $A, B \geq V/(4\pi D) > 0$. So we do not have arbitrarily thin tori with $B_j \rightarrow 0$ in this class.

Once Cheeger had a uniform lower bound on injectivity radius, he proved

$$\mathcal{M}_{K_1, V}^{K_2, D} \subset \mathcal{M}_{K_1 i_0}^{K_2, D} \quad (4.28)$$

where

$$\mathcal{M}_{K_1 i_0}^{K_2, D} = \{M^n \mid K_1 \leq \text{sect} \leq K_2, \text{in}jrad \geq i_0, \text{Diam}(M) \leq D\}. \quad (4.29)$$

Note that any manifold M^n in this second family can be covered by at most $N_{K_1 i_0}^{K_2, D}$ charts as follows. About any point $p \in M^n$, we can take a chart defined by the exponential map, \exp_p , which covers a ball about p of radius $r = i_0$. To choose a finite collection, we take a maximal collection of disjoint balls of radius $r/2$, so that the balls of radius r cover M^n . Let

$$N(M, r) = \text{max number of disjoint balls of radius } r/2 \text{ in } M. \quad (4.30)$$

To find a uniform bound on this number, note that $\text{Vol}(B(p, r/2)) \geq V_{K_2}^n(r/2)$ which is the volume of a ball in a simply connected space of constant sectional curvature K_2 , so

$$N(M, r) V_{K_2}^n(r/2) \leq \sum_{i=1}^{N(M, r)} \text{Vol}(B(p_i, r/2)) \leq \text{Vol}(M) \leq V_{K_1}^n(D). \quad (4.31)$$

Thus

$$N(M, r) \leq N(r) = V_{K_1}^n(D) / V_{K_2}^n(r/2). \quad (4.32)$$

Note that the lower bound on the $\text{Vol}(B(p, r/2))$ here strongly uses the fact that $r \leq i_0 \leq \text{injr}(p)$. Clearly such a uniform lower bound on volume does not hold on increasingly thin tori with $B_j \rightarrow 0$.

Once we have a uniform bound on the number of charts, we see that we have only finitely many homeomorphism types bounded by all the possible ways these charts can overlap. If we have a sequence of M_j in a class \mathcal{M} with a uniform bound $N(M, r) \leq N(r)$ then we can apply the pigeon hole principle to choose a subsequence, M_j , such that they are all homeomorphic and such that $N(M_j, r)$ is constant and such that the same pairs of balls of radius r intersect. To complete the proof of Cheeger's Compactness Theorem, we would then only need to show uniform bounds on the sequence of metric tensors within these charts. We recommend Petersen's textbook for a discussion of the many ways that metric tensors can be bounded within the charts [79]. See also Gromov's textbook [55].

4.7 Injectivity Radius, Heat Kernels, and Eigenvalues

We have seen above in Sections 2.5 and 3.10 there are entire classes of manifolds which share the same bounds on their heat kernels and on their eigenvalues. In particular there was the class

$$\mathcal{M}_H^n, i_0 = \{M^n \mid \text{Ricci}_M \geq (n-1)H \text{ and } \text{injr}_M \geq i_0\}. \quad (4.33)$$

The injectivity radius bound here is essential. The Ricci curvature bound is also a key step.

Consider for example the rescaled circles \mathbb{S}_R^1 that we studied in Section 2.2 and 3.5. Their injectivity radii

$$\text{injr}_{\mathbb{S}_R^1} = \pi R \rightarrow 0 \text{ as } R \rightarrow 0. \quad (4.34)$$

We saw in Section 2.2 that their heat kernels,

$$h_t^R(\theta_1, \theta_2) = R^{-1} h_{t/R^2}(\theta_1, \theta_2) \tag{4.35}$$

are not uniformly bounded above for sequences as $R \rightarrow 0$. We saw in Section 3.5 that their eigenvalues diverge

$$\lambda_{2j} = \lambda_{2j-1} = j^2/R^2 \rightarrow \infty \text{ as } R \rightarrow 0. \tag{4.36}$$

Sequences of rescaled spheres, \mathbb{S}_R^2 , have nonnegative Ricci curvature and the same problems with their injectivity radii, heat kernels, and eigenvalues as the \mathbb{S}_R^1 as $R \rightarrow 0$.

4.8 Harmonic Coordinates and Harmonic Radius

In [43], DeTurck and Kazdan proved that whenever an atlas exists covering M with charts such that $h \circ \varphi_i \subset C^k(U_i)$, then the same is true if we chose the atlas of **harmonic coordinate charts**. A coordinate chart $\varphi : U \subset \mathbb{R}^m \rightarrow M$ is **harmonic** if the components of φ^{-1} are harmonic functions:

$$\varphi^{-1}(p) = (x_1(p), x_2(p), \dots, x_n(p)) \text{ where } \Delta x_i = 0. \tag{4.37}$$

A manifold M has **harmonic radius** $r_h > 0$ if it can be covered by a collection of harmonic coordinate charts, each of which covers a ball of radius r_h .

In 1990, Anderson [7] proved a $C^{1,\alpha}$ precompactness theorem for

$$\mathcal{M}_{H,i_0}^D = \{M^n : |\text{Ricci}| \leq (n - 1)H, \text{ injrad}(M) \geq i_0 > 0, \text{ Diam}(M) \leq D\} \tag{4.38}$$

using harmonic coordinate charts by finding a uniform lower bound on the harmonic radius, which is a lower bound on the radius of harmonic coordinate charts. Building upon this work, Anderson-Cheeger proved a C^α compactness theorem for a larger class of manifolds in [8].

The harmonic radius can also be used to control eigenfunctions as is seen for example in the work of Portegies [81] that we will present at the end of this paper.

4.9 Gromov-Hausdorff Convergence

In 1983 Gromov introduced the notion of the Gromov-Hausdorff distance between compact metric spaces [55]. Since a compact Riemannian manifold, (M, g) can be viewed as metric space, (M, d_g) , his notion allowed him to define

the Gromov-Hausdorff convergence of Riemannian manifolds to limits which may only be metric spaces. The notion allowed him to study sequences of collapsing manifolds (like increasingly thin tori) whose limits have lower dimension as well as sequences of noncollapsing manifolds.

Before we define Gromov-Hausdorff distance we review that the **Hausdorff distance** between two compact sets, K_1 and K_2 , in a metric space, Z , is defined

$$d_H^Z(K_1, K_2) = \inf\{R : K_1 \subset T_R(K_2) \text{ and } K_2 \subset T_R(K_1)\} \quad (4.39)$$

where $T_R(K_i)$ is the tubular neighborhood about K_i :

$$T_R(K_i) = \{z \in Z : \exists x \in K_i \text{ s.t. } d_Z(x, z) < R\}. \quad (4.40)$$

This is an extrinsic way of measuring the distance between K_1 and K_2 which depends very much on the extrinsic space, Z , in which they both sit.

To define an intrinsic distance between two compact metric spaces, (X_1, d_1) and (X_2, d_2) , which are not sitting in a common metric space, Gromov considered all distance preserving maps,

$$F_i : X_i \rightarrow Z \text{ such that } d_Z(F_i(p), F_i(q)) = d_{X_i}(p, q) \quad \forall p, q \in X_i \quad (4.41)$$

and all possible compact metric spaces, (Z, d_Z) . He defined the **Gromov-Hausdorff distance** by taking the infimum over all such possibilities:

$$d_{GH}(X_1, X_2) = \inf d_H^Z(F_1(X_1), F_2(X_2)). \quad (4.42)$$

He called this an intrinsic Hausdorff distance because it depends only on intrinsic information about the pair of metric spaces, (X_1, d_1) and (X_2, d_2) , and is defined using the Hausdorff distance.

Gromov proved that if a sequence $(X_j, d_j) \rightarrow (X_\infty, d_\infty)$ in the Gromov-Hausdorff sense then there exists a common compact metric space, Z , and distance preserving maps $F_j : X_j \rightarrow Z$ such that

$$d_H^Z(F_j(X_j), F_\infty(X_\infty)) \rightarrow 0. \quad (4.43)$$

In particular, for all $p_\infty \in X_\infty$ there exists $p_j \in X_j$ such that $F_j(p_j) \rightarrow F_\infty(p_\infty)$. Although this is not uniquely assigned, this is how

$$p_j \in X_j \text{ converges to } p_\infty \in X_\infty \quad (4.44)$$

is defined.

Gromov's Compactness Theorem: *A sequence of Riemannian manifolds*

$$M_j^n \in \mathcal{M}_{H,D}^n = \{M^n \mid \text{Ricci}_{M^n} \geq (n-1)H, \text{Diam}(M^n) \leq D\}. \quad (4.45)$$

has a subsequence which converges in the Gromov-Hausdorff sense to a compact metric space.

Gromov proved this in [55] by first proving the following theorem:

Bishop-Gromov Volume Comparison Theorem: *for any $M^n \in \mathcal{M}_{H,D}^n$, and any $p \in M^n$ and any $R > r > 0$, the volumes of balls satisfy the following:*

$$\text{Vol}(B(p, r)) / \text{Vol}(B(p, R)) \geq V_H^n(r) / V_H^n(R) \quad (4.46)$$

where $V_H^n(s)$ is the volume of a ball of radius s in an n dimensional simply connected manifold of constant sectional curvature, H .

This allowed Gromov to uniformly count the number, $N(r)$, of balls of any given radius in $M_j^n \in \mathcal{M}_{H,D}^n$ space. He applied this to create uniformly dense countable collection points, $p_{j,i} \in M_j$ and matrices of distances between these points,

$$d_{j,i,k} = d_{M_j}(p_i, p_k) \in [0, D]. \quad (4.47)$$

For each, (i, k) , a subsequence of j (that we still denote by j) has

$$d_{j,i,k} \rightarrow d_{\infty,j,k} \in [0, D]. \quad (4.48)$$

Gromov diagonalized the sequence to obtain this convergence for all pairs of points and then built a limit space out of the countable collection of points assigning them

$$X = \{p_j : j \in \mathbb{N}\} \text{ and } d_{\infty}(p_j, p_k) = d_{\infty,j,k}. \quad (4.49)$$

He used the triangle inequality and the uniform density of the points to take a metric completion, \bar{X} , and prove $M_j \rightarrow \bar{X}$ in the Gromov-Hausdorff sense for this subsequence. See [55] and see also the text of Burago-Burago-Ivanov [18]).

4.10 Increasingly Thin Tori GH-Converge to Circles

An example of a collapsing sequence of $M_j \subset \mathcal{M}_{H,D}^n$ which converges in the Gromov-Hausdorff sense is the sequence of increasingly thin tori

$$\mathbb{S}_A^1 \times \mathbb{S}_{1/j}^1 \xrightarrow{\text{GH}} \mathbb{S}_A^1. \quad (4.50)$$

This can be seen by observing that for any fixed j , we can take

$$Z = \mathbb{S}_A^1 \times \mathbb{S}_{1/j}^1 \quad (4.51)$$

and distance preserving maps

$$F_1 : \mathbb{S}_A^1 \times \mathbb{S}_{1/j}^1 \rightarrow Z \text{ to be } F_2(\theta, \varphi) = (\theta, \varphi). \quad (4.52)$$

and

$$F_2 : \mathbb{S}_A^1 \rightarrow Z \text{ to be } F_2(\theta) = (\theta, 0). \quad (4.53)$$

The Hausdorff distance between

$$F_2(\mathbb{S}_A^1) = \{(\theta, 0) : \theta \in [0, 2\pi]\} \subset Z \text{ and } F_1(\mathbb{S}_A^1 \times \mathbb{S}_{1/j}^1) = Z \quad (4.54)$$

is easily seen to be π/j . Thus

$$d_{GH}(\mathbb{S}_A^1 \times \mathbb{S}_{1/j}^1, \mathbb{S}_A^1) < \pi/j \rightarrow 0. \quad (4.55)$$

In Section 3.7, we saw that for each i the i^{th} eigenfunctions of increasingly thin tori $\mathbb{S}_A^1 \times \mathbb{S}_B^1$ converge to the i^{th} eigenfunctions of the circle, \mathbb{S}_A^1 . Thus the truncated heat kernels of thin tori $\mathbb{S}_A^1 \times \mathbb{S}_B^1$ rescaled by $(2\pi B)$ converge as $B \rightarrow 0$ to

$$\frac{1}{2\pi} + \frac{1}{\pi} \sum_{j=1}^N e^{-j^2 t} (\sin(j\theta_1) \sin(j\theta_2) + \cos(j\theta_1) \cos(j\theta_2))$$

which is the truncated heat kernel on a circle \mathbb{S}_A^1 . Indeed the full heat kernel of $\mathbb{S}_A^1 \times \mathbb{S}_B^1$ rescaled by $(2\pi B)$ converges to the heat kernel of \mathbb{S}_A^1 as $B \rightarrow 0$.

4.11 Eigenvalues and Measured Gromov-Hausdorff Convergence

This nice behavior of the eigenvalues and heat kernels is not true in general. Fukaya found examples of sequences of Riemannian manifolds which converge in the Gromov-Hausdorff sense to lower dimensional manifolds such that the eigenfunctions did not converge the eigenvalues of the limit manifold [53]. More precisely, the eigenvalues, $\lambda_k(M_j)$, do not converge to eigenvalues, $\lambda_k(M_\infty)$ of the standard Laplace operator, Δ_{M_∞} , on the limit manifold, M_∞ .

Fukaya studied collapsing warped tori:

$$(M_j, g_j) = (\mathbb{S}^1 \times \mathbb{S}^1, d\theta^2 + f^2(\theta)/j^2 d\phi^2) \quad (4.56)$$

where $f(\theta)$ is a nonconstant positive periodic warping function, which converge in the Gromov-Hausdorff sense to the standard $(\mathbb{S}^1, d\theta^2)$. Since the Laplacian on (M_j, g_j) has the form

$$\Delta_{M_j} = 1 \frac{1}{f(\theta)} \frac{\partial}{\partial \theta} f(\theta) \frac{\partial}{\partial \theta} - \frac{1}{j^2 f^2(\theta)} \frac{\partial^2}{\partial \phi^2}. \quad (4.57)$$

Fukaya explained that their eigenvalues, $\lambda_k = \lambda_k(M_j)$ converge as $j \rightarrow \infty$, to the eigenvalues, λ'_k , of the operator

$$1 \frac{1}{f(\theta)} \frac{d}{d\theta} f(\theta) \frac{d}{d\theta} \tag{4.58}$$

on \mathbb{S}^1 . This operator is not equal to the Laplace operator, $\Delta_{\mathbb{S}^1} = d^2/d\theta^2$, and has different eigenvalues, $\lambda'_k \neq \lambda_k(\mathbb{S}^1)$.

In [53], Fukaya proposed the notion of **measured Gromov-Hausdorff convergence** of metric measure spaces:

$$(X_j, d_j, \mu_j) \rightarrow (X_\infty, d_\infty, \mu_\infty) \tag{4.59}$$

where $(X_j, d_j) \rightarrow (X_\infty, d_\infty)$ in the Gromov-Hausdorff sense and if $p_j \in X_j$ converges to $p_\infty \in X_\infty$ as in (4.44) then the volumes of balls about them converge:

$$\mu_j(B(p_j, r)) \rightarrow \mu_\infty(B(p_\infty, r)) \text{ for all } r > 0. \tag{4.60}$$

Fukaya proved that with this notion of metric measure convergence, and a natural kind of Laplacian defined in a way that depends on the measure, there is semicontinuity of the eigenvalues:

$$\limsup_{j \rightarrow \infty} \lambda_k(X_j, d_j, \mu_j) \leq \lambda_k(X_\infty, d_\infty, \mu_\infty). \tag{4.61}$$

Fukaya conjectured that eigenvalues and other spectral properties will behave well under metric measure convergence if the sequence is in $\mathcal{M}_{H,D}^n$ and he proved this conjecture in the case with two sided sectional curvature bounds. He found counter examples without the lower bound on Ricci curvature.

In [25], Cheeger and Colding proved Fukaya's conjecture for sequences of $M_j \in \mathcal{M}_{H,D}^n$. In fact, they proved that any sequence $M_j \in \mathcal{M}_{H,D}^n$ endowed with measures defined using rescaled volumes has a subsequence which converges in the metric measure sense to a metric measure space with a measure that satisfies the Bishop-Gromov Inequality. They defined a natural notion of eigenfunction on this limit space and proved the eigenfunctions converge as well. Building on their work, Yu Ding proved the heat kernels converge as well [45].

More recently Sturm has defined the metric measure distance between Riemannian manifolds is defined as the infimum over all compact metric spaces Z and all distance preserving maps $F_i : M_i \rightarrow Z$ of the Wasserstein distance between the push forwards of their measures into Z [90]. This notion is useful for studying $CD(H, n)$ spaces. These $CD(H, n)$ spaces were first introduced by

Lott and Villani and by Sturm using optimal transport [72][90] and were shown to include the limits of $M_j \in \mathcal{M}_{H,D}^n$. See also Villani’s text [94]. A stronger class of spaces called $\text{RCD}^*(H, n)$ spaces which were introduced by Ambrosio-Gigli-Savare in [5] have all the properties of $CD(H, n)$ spaces and also satisfy a splitting theorem. They proved all limits of manifolds $M_j \in \mathcal{M}_{H,D}^n$ are included in this stronger class as well.

It is also of interest to study sequences of manifolds which do not have uniform bounds on their Ricci or sectional curvatures.

4.12 Barbells

In [53], Fukaya considered sequences of **barbells** which are pairs of spheres of radius R_1 and R_2 attached by a cylinders of length L and radius $r \leq \min\{R_1, R_2\}$ that can be described by as warped products over a sphere:

$$(M_{R_1, R_2, L, r}, g_{R_1, R_2, L, r}) = (\mathbb{S}^2, ds^2 + f^2(s)d\theta^2) \tag{4.62}$$

where

$$\begin{aligned} f(s) &= R_1 \sin((s + s_{R_1})/R_1) \text{ for } s \in [-s_{R_1}, -L/2] \\ f(s) &= r \text{ for } s \in [-L/2, L/2] \\ f(s) &= R_2 \sin((s_{R_2} - s)/R_2) \text{ for } s \in [L/2, s_{R_2}] \end{aligned} \tag{4.63}$$

where s_{R_i} are chosen so that we have a C^0 metric:

$$R_1 \sin((L/2 + s_{R_1})/R_1) = r = R_2 \sin((s_{R_2} - L/2)/R_2). \tag{4.64}$$

These are C^0 Riemannian manifolds, but can easily be smoothed to be C^∞ Riemannian manifolds with metric tensors arbitrarily close in the C^0 sense to the given metric tensor in (4.63). However the Ricci curvature of the sequence approaching this decreases to negative infinity for any such sequence because these spaces do not satisfy the Bishop-Gromov inequality.

Fukaya fixed $R_2 = R_1$ and fixed $L > 0$ and took $r \rightarrow 0$. He explained that the sequence of such barbells converges in the measured Gromov-Haudorff sense to a metric measure space which is a pair of standard spheres joined by a line segment of length L :

$$M_{R_1, R_2, L, r} \xrightarrow{\text{mGH}} X_{R_1, R_2, L} = \mathbb{S}_{R_1}^1 \sqcup | \sim [0, L] \sqcup | \sim \mathbb{S}_{R_2}^1 \text{ as } r \rightarrow 0 \tag{4.65}$$

where one pole \mathbb{S}_{R_1} is identified to $0 \in [0, L]$ and one pole in \mathbb{S}_{R_2} is identified to $L \in [0, L]$. See Figure 4.1.

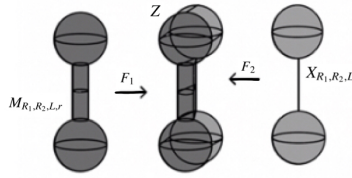


Figure 4.1: Here we see distance preserving maps $F_1 : M_{R_1, R_2, L, r} \rightarrow Z$ and $F_2 : X_{R_1, R_2, L} \rightarrow Z$ that can be used to estimate the Gromov-Hausdorff distance, $d_{GH}(M_{R_1, R_2, L, r}, X_{R_1, R_2, L}) \approx \pi r$.

Fukaya noted measure on this limit space is the standard measure on each sphere and is 0 on the interval. He also studied sequences where the metric on the cylinder was increasingly bumpy so its area would not converge to 0 and the measure on the interval would be nonzero. Furthermore he proved he had strong enough control on this sequence, that the eigenvalues of the limit space are found by taking all the eigenvalues of $\mathbb{S}_{R_1}^1$ and all the eigenvalues of $\mathbb{S}_{R_2}^1$ and all the eigenvalues of the interval $[0, L]$ and listing them in order with multiplicity. Despite the lack of uniform control on Ricci curvature, he stated that the eigenvalues of the barbells converge to these eigenvalues. It would be interesting to investigate what the eigenfunctions and heat kernels look like on this sequence.

Fukaya also studied that case where R_1 is fixed and $L \rightarrow 0$ and $r \rightarrow 0$ and $R_2 > r \rightarrow 0$ so that

$$M_{R_1, R_2, r, L} \xrightarrow{\text{GH}} \mathbb{S}_{R_1}^2 \text{ as } r \rightarrow 0, L \rightarrow 0, R_2 \rightarrow 0. \tag{4.66}$$

By taking $r \rightarrow 0$ faster than the other two, he can use his previous result to estimate the eigenvalues. In particular for r sufficiently small,

$$M_{R_1, R_2, r, L} \approx X_{R_1, R_2, L} \tag{4.67}$$

and then

$$X_{R_1, R_2, L} \xrightarrow{\text{GH}} \mathbb{S}_{R_1}^2 \text{ as } r \rightarrow 0, L \rightarrow 0, R_2 \rightarrow 0. \tag{4.68}$$

He studied the eigenvalues in this setting and it would be interesting to study the eigenfunctions and heat kernels as well.

4.13 Intrinsic Flat Convergence

The Sormani-Wenger intrinsic flat distance between a pair of oriented Riemannian manifolds, (M_1^n, g_1) and (M_2^n, g_2) , is essentially the filling volume between them. It is defined by taking the infimum over all complete metric spaces Z and all distance preserving maps, $F_i : (M_i, d_{g_i}) \rightarrow (Z, d_z)$ of the flat distance between their images:

$$d_{SWIF}(M_1, M_2) = \inf d_F^Z(F_{1\#}[[M_1]], F_{2\#}[[M_2]]). \quad (4.69)$$

As it requires Geometric Measure Theory and Ambrosio-Kirchheim Theory to define the flat distance between two submanifolds, we instead state a simple useful estimate:

$$d_F^Z(F_{1\#}[[M_1]], F_{2\#}[[M_2]]) \leq \text{Vol}_n(A^n) + \text{Vol}_{n+1}(B^{n+1}) \quad (4.70)$$

where A^n is an n -dimensional submanifold of Z and the filling manifold, B^{n+1} , is an $(n + 1)$ -dimensional submanifold of Z such that the boundary of B is $F_{1\#}[[M_1]], F_{2\#}[[M_2]]$ and A keeping track of the orientation:

$$\partial B = F_{1\#}[[M_1]] - F_{2\#}[[M_2]] - A. \quad (4.71)$$

More generally, A and B are integral currents satisfying (4.71) and their volumes are replaced by their Ambrosio-Kirchheim masses. In fact the intrinsic flat distance is defined between pairs of **integral current spaces** which are rectifiable metric spaces with rectifiable boundaries that have a weighted orientation called an integral current structure. The intrinsic flat limits of sequences of oriented Riemannian manifolds are integral current spaces included possibly the 0 space when the sequence collapses [89].

This notion of convergence was defined to study the noncollapsed limits of sequences of Riemannian manifolds with nonnegative scalar curvature. It is designed so that thin wells disappear in the limit.

Sormani-Wenger studied spheres with wells that our student research team liked to call **lollipops**. These consist of a single sphere of radius R attached to a capped cylinder of length L and radius $r \leq R$ in [89]. These manifolds

$$(M_{R,r,L}, g_{R,r,L}) = (\mathbb{S}^2, ds^2 + f^2(s)d\theta^2) \quad (4.72)$$

are defined using the following C^0 warping functions:

$$\begin{aligned} f(s) &= R \sin((s + s_R)/R) \text{ for } s \in [-s_R, -L/2] \\ f(s) &= r \text{ for } s \in [-L/2, L/2] \\ f(s) &= r \sin((s_r - s)/R_2) \text{ for } s \in [L/2, s_r] \end{aligned}$$

as in (4.63) with $R_1 = R$ and $R_2 = r$. If we fix $R = r > 0$ and take $L \rightarrow 0$ these converge in the C^0 sense to a rescaled sphere of radius R :

$$M_{R,r,L} \rightarrow \mathbb{S}_R^2 \text{ as } L \rightarrow 0. \tag{4.73}$$

If we fix $R, L > 0$ and take $r_j \rightarrow 0$, these have no C^0 limit, but converge in the mGH sense to a sphere of radius R attached to an interval of length L :

$$M_{R,r_j,L} \xrightarrow{mGH} X_{R,L} = \mathbb{S}^1 \sqcup [0, L] \text{ as } r_j \rightarrow 0. \tag{4.74}$$

See Figure 4.2. The measure on this limit space is the standard measure on the sphere and is 0 on the interval. Sormani-Wenger showed intrinsic flat limit is just the sphere by explicitly constructing distance preserving maps into a sequence of complete spaces Z_j and constructing fillings B_j of one dimension higher between the sphere and the lollipop that has very small volume. Portegies studied the eigenvalues of this example in [80] and it would be interesting to know how the eigenfunctions and heat kernels behave.

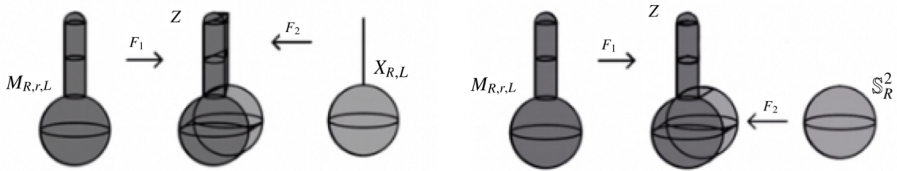


Figure 4.2: On the left we see distance preserving maps $F_1 : M_{R,L,r} \rightarrow Z$ and $F_2 : X_{R,L} \rightarrow Z$ that can be used to estimate the Gromov-Hausdorff distance, $d_{GH}(M_{R,L,r}, X_{R,L}) \approx \pi r$. On the right we see the same maps can be used to show the intrinsic flat distance between $M_{R,L,r}$ and \mathbb{S}_R^2 is small.

If the warping functions on the sequence are increasingly bumpy on the cylindrical part instead of being constant, Sormani-Wenger showed the measured Gromov-Hausdorff limit has a nonzero measure on the interval but the intrinsic flat limit is still only the sphere [89].

Ilmanen created a sequence of spheres with increasingly many increasingly thin capped cylinders with no Gromov-Hausdorff limit because such a sequence cannot be embedded with distance preserving maps into a compact metric space, Z . Sormani-Wenger showed that if the volume of the sequence is kept bounded then the filling volumes between these manifolds and the sphere converges to 0,

so that the intrinsic flat limit of Ilmanen's Example is a standard sphere. The Intrinsic Flat convergence was designed to handle sequences like these that have no Gromov-Hausdorff limit. Wenger's Compactness Theorem states that any sequence of manifolds with $\text{Vol}(M_j) \leq V$ and $\text{Diam}(M_j) \leq D$ has a subsequence which converges in the intrinsic flat sense to an integral current space.

In [80], Portegies introduced **volume preserving intrinsic flat convergence**, $M_j \rightarrow M_\infty$ iff

$$M_j \rightarrow M_\infty \text{ in the intrinsic flat sense and } \text{Vol}(M_j) \rightarrow \text{Vol}(M_\infty). \quad (4.75)$$

Applying the work of Fukaya, he proved that under this convergence

$$\limsup_{j \rightarrow \infty} \lambda_k(M_j) \leq \lambda_k(M_\infty). \quad (4.76)$$

It would be interesting to study what happens to the heat kernels in this setting as well.

4.14 VADB Convergence

Most recently the notion of VADB Convergence has been defined by Allen-Perales-Sormani in [2] and they proved VADB convergence implies Intrinsic Flat Convergence. Ilmanen's Example converges to a sphere under VADB convergence.

A sequence of diffeomorphic manifolds (M, g_j) converge to (M, g_0) in the **VADB** sense if $\text{Vol}_{g_j}(M) \rightarrow \text{Vol}_{g_0}(M)$, $\text{Diam}_{g_j}(M) \leq D$, and $g_j \geq (1 - 1/j)g_0$.

This paper is building on prior work of Allen-Sormani concerning the L^p convergence of the metric tensors which is not enough to achieve uniform, Gromov-Hausdorff, or intrinsic flat convergence unless there are also uniform bounds on the metric tensors from above and below as shown by Allen-Sormani in [3] and [4].

A number of conjectures concerning VADB convergence as well as a survey of results applying VADB convergence appears in a survey by the second author [88]. It would be interesting to study how the eigenvalues, eigenfunctions, and heat kernels behave under VADB convergence.

5 Embedding Riemannian Manifolds via the Heat Kernel

In 1994 Bérard, Besson, and Gallot introduced the idea of embedding a Riemannian manifold into ℓ^2 using the heat kernel [16]. Recall that ℓ^2 is the vector

space of square summable sequences,

$$\ell^2 = \left\{ \{v_j\}_{j=0}^\infty = \{v_0, v_1, v_2, v_3, \dots\} : \sum_{i=0}^\infty v_i^2 < \infty \right\}, \quad (5.1)$$

with the ℓ^2 -inner product and ℓ^2 -norm,

$$\langle \{v_j\}_{j=0}^\infty, \{\tilde{v}_j\}_{j=0}^\infty \rangle_{\ell^2} = \sum_{j=0}^\infty v_j \cdot \tilde{v}_j \quad \text{and} \quad \|\{v_j\}_{j=0}^\infty\|_{\ell^2}^2 = \sum_{j=0}^\infty v_j^2. \quad (5.2)$$

Given an $L^2(M)$ orthonormal basis of eigenfunctions

$$\{\phi_0, \phi_1, \phi_2, \dots\} \text{ such that } \Delta\phi_j = -\lambda_j\phi_j \text{ with } 0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots, \quad (5.3)$$

Bérard-Besson-Gallot define the family of maps $\Psi_t : M \rightarrow \ell^2$ for $t > 0$ as

$$\Psi_t(x) = \left\{ e^{-\lambda_j t/2} \phi_j(x) \right\}_{j=1}^\infty. \quad (5.4)$$

They use the fact that

$$\begin{aligned} \|\Psi_t(x) - \Psi_t(y)\|_{\ell^2}^2 &= \sum_{j=1}^\infty e^{-\lambda_j t} (\phi_j(x) - \phi_j(y))^2 = \sum_{j=0}^\infty e^{-\lambda_j t} (\phi_j(x) - \phi_j(y))^2 \\ &= \sum_{j=0}^\infty e^{-\lambda_j t} (\phi_j(x))^2 + \sum_{j=0}^\infty e^{-\lambda_j t} (\phi_j(y))^2 - 2 \sum_{j=0}^\infty e^{-\lambda_j t} \phi_j(x) \phi_j(y) \\ &= h_t(x, x) + h_t(y, y) - 2h_t(x, y) \end{aligned}$$

to prove the map is continuous and thus the image is compact.

To prove this is an embedding, Bérard-Besson-Gallot use the fact that any $L^2(M)$ basis of orthonormal eigenfunctions *separates points*:

$$\forall p, q \in M \exists \phi_j \text{ such that } \phi_j(p) \neq \phi_j(q). \quad (5.5)$$

This is also used to prove that the $L^2(M)$ basis of orthonormal eigenfunctions spans $L^2(M)$.

Note that the the eigenvalues, $0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \dots$ and the eigenspaces of these eigenvalues are uniquely determined by M and these eigenspaces are $L^2(M)$ -orthogonal to one another. There is a choice being made when selecting the $L^2(M)$ -orthonormal basis for each eigenspace, and thus the space $L^2(M)$ -orthonormal eigenfunctions depend on this choice. For this reason Bérard, Besson and Gallot label their eigenfunctions with an a .

5.1 Rescaled Heat Kernel Embeddings

Bérard, Besson, and Gallot [16] defined rescaled heat kernel embeddings and proved the following theorem:

Theorem 5.1. Given an n -dimensional closed Riemannian manifold, let the map $\psi_t : M^n \rightarrow \ell^2$ for $t > 0$ be defined by

$$\psi_t(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \Psi_t(x). \quad (5.6)$$

Then the pulled back metric tensor satisfies

$$(\psi_t)^* g_{\ell^2} = g + \frac{t}{3} \left(\frac{1}{2} \text{Scal}_g \cdot g - \text{Ricci}_g \right) + O(t^2) \text{ when } t \rightarrow 0^+ \quad (5.7)$$

where the pulled back metric tensor is defined on a vector $V_x \in TM_x$ by

$$(\psi_t)^* g_{\ell^2}(V_x, V_x) = \|d\psi_t V_x\|_{\ell^2}^2 = \left(\sqrt{2}(4\pi)^{n/2} t^{(n+2)/4} \right)^2 \sum_{j=1}^{\infty} e^{-\lambda_j t} |d\phi_j V_x|^2. \quad (5.8)$$

Bérard-Besson-Gallot proved (5.7) by observing that

$$\sum_{j=1}^{\infty} e^{-\lambda_j t} |d\phi_j V_x|^2 = \partial_{y_1} \partial_{y_2} h_t(y_1, y_2)|_{y_1=y_2=x}. \quad (5.9)$$

They were able to prove their theorem taking these partial derivatives on the heat kernel using the Minakshisundaram-Pleijel expansion.

Minakshisundaram and Pleijel studied the heat kernel on a manifold M within a normal neighborhood about a point x using the method of analytic continuation applied to the time variable [77]. They required that the metric tensor g of M is an analytic function of the normal coordinates. They prove the heat kernel is the C^ω limit

$$h_t(x, y) = \lim_{P \rightarrow \infty} \tilde{h}_t^P(x, y) \quad (5.10)$$

where

$$\tilde{h}_t^P(x, y) = \frac{1}{(4\pi t)^{n/2}} e^{-d_M(x,y)^2/(4t)} \sum_{i=0}^P U_i(x, y) t^i \quad (5.11)$$

where $U_i(x, y)$ satisfy an iterative set of partial differential equations (each depending on the value of the previous $U_i(x, y)$ and the metric tensor g). They proved the $U_i(x, y)$ are uniquely determined by the requirement that $U_0(x, x) = 1$ and $U_i(x, x)$ is finite. One advantage of Minakshisundaram-Pleijel's expansion

is that a C^ω limit allows us to differentiate as many times as we wish and the derivatives will converge to the derivatives of the limit. The proof of convergence in [77] involves the decay of the eigenvalues. The disadvantage of this formula is that it requires that x, y lie in a normal neighborhood.

Abdalla later applied Minakshisundaram and Pleijel to study C^ω parametrized families of Riemannian manifolds embedded via heat kernel maps in [1].

5.2 Truncated Heat Kernel Embeddings

It is much simpler to consider the truncated heat kernel embeddings into Euclidean space $\psi_t^N : M^n \rightarrow \mathbb{E}^N$ for $t > 0$ such that

$$\psi_t^N(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \{e^{-\lambda_j t/2} \phi_j(x)\}_{j=1}^N \in \mathbb{E}^N. \quad (5.12)$$

Note that we easily obtain continuity for the truncated heat kernel maps and they are one-to-one as long as $\{\phi_1, \dots, \phi_N\}$ separate points.

We can then rewrite Bérard-Besson-Gallot's Theorem stated in (5.7) as

$$\lim_{t \rightarrow 0^+} \lim_{N \rightarrow \infty} (\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) = g(V_x, V_x), \quad (5.13)$$

and

$$\lim_{t \rightarrow 0^+} (3/t) \left(\lim_{N \rightarrow \infty} (\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) - g(V_x, V_x) \right) = \frac{1}{2} \text{Scal}_g \cdot g - \text{Ricci}_g \quad (5.14)$$

where

$$(\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) = \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 \sum_{j=1}^N e^{-\lambda_j t} |d\phi_j V_x|^2. \quad (5.15)$$

Note that these limits must be taken in the correct order. If we try to take the limit as $t \rightarrow 0^+$ first in (5.13), then

$$\lim_{t \rightarrow 0^+} (\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) = \lim_{t \rightarrow 0^+} \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 \sum_{j=1}^N e^{-\lambda_j t} |d\phi_j V_x|^2 = 0 \quad (5.16)$$

because $\lim_{t \rightarrow 0^+} t^{(n+2)/4} e^{-\lambda t} = 0$.

We can define a BBG time T_ϵ such that

$$|(\psi_t)^* g_{\ell^2}(V_x, V_x) - g(V_x, V_x)| < \epsilon g(V_x, V_x) \quad \forall t \in (0, T_\epsilon) \quad (5.17)$$

and a BBG number $N_{t,\varepsilon}$ such that

$$|(\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) - (\psi_t)^* g_{\ell^2}(V_x, V_x)| < \varepsilon g(V_x, V_x) \quad \forall N \geq N_{t,\varepsilon} \quad (5.18)$$

If we choose $t < T_\varepsilon$ first and then $N > N_{t,\varepsilon}$ we have

$$|(\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) - g(V_x, V_x)| < (\varepsilon + \varepsilon)g(V_x, V_x). \quad (5.19)$$

These estimates are explored further in the work of Portegies [81] which we will discuss below in Section 7.4.

Note that there is some lack of uniqueness when defining the truncated embeddings. If N has $\lambda_N = \lambda_{N+1}$, then ψ_t^N includes ϕ_N in the truncated embedding but doesn't include ϕ_{N+1} although both are in the same eigenspace and could have been listed in any order when first finding the $L^2(M)$ basis of eigenfunctions. If N has $\lambda_N < \lambda_{N+1}$, then the truncated embedding is unique up to an orthogonal transformation of \mathbb{E}^N , because we would be including a complete $L^2(M)$ basis for each eigenspace with $j \leq N$ and choosing a different basis for the m_j dimensional eigenspace of λ_j only transforms the basis by an $O(m_j)$ action and thus transforms the image in \mathbb{E}^N by the same action.

5.3 Embedding Circles with a Truncated Heat Kernel

Let us consider $\psi_t : \mathbb{S}^1 \rightarrow \ell^2$ for $t > 0$ using the eigenfunctions (which come in pairs) that we found in the Section 3.4. The truncated embedding is

$$\psi_t^{2N}(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \left\{ e^{-j^2 t/2} \pi^{-\frac{1}{2}} \sin(j\theta), e^{-j^2 t/2} \pi^{-\frac{1}{2}} \cos(j\theta) \right\}_{j=1}^N \subset \mathbb{E}^{2N}.$$

In Figure 5.1, we have drawn some of the images of the truncated heat kernel embeddings with $T > t$. Note that already with $N = 2$ we have an embedding, but the circle is too short. With $N = 4$ we wind around $\psi_T^2(\mathbb{S}^1)$ twice, and with $N = 6$ we wind around $\psi_T^4(\mathbb{S}^1)$ thrice. Taking limit as $N \rightarrow \infty$ for fixed time T , we moving to the right across the first row of Figure 5.1. The limit, $\Psi_T(\mathbb{S}^1)$ is a circle winding infinitely many times around itself. The same happens in the second row.

Now consider the columns in Figure 5.1. If we take t small, for fixed N , we see that the radius of winding is closer to the original radius, because $e^{-j^2 t}$ is close to 1 for both $j = 1$ and $j = 2$. However, the whole image is scaled down by $\sqrt{2}(4\pi)^{n/2} t^{(n+2)/4}$. So for fixed N taking $t \rightarrow 0$ the images contract to a point.

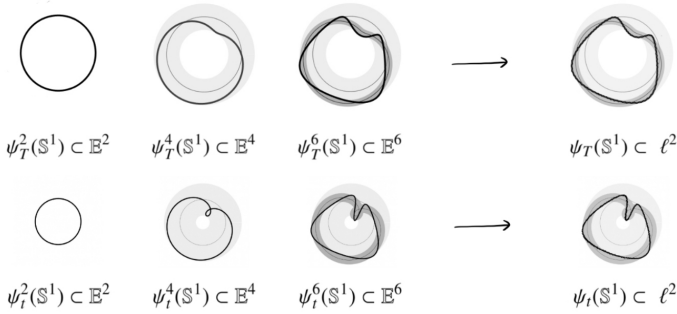


Figure 5.1: Here we look at the images of the heat kernel embeddings of the circle for $T > t$. Image credit: Pen Chang (Student, Stony Brook).

To find the lengths of these curves, we use the homogeneity of \mathbb{S}^1 to see that

$$L(\psi_t^{2N}(\mathbb{S}^1)) = \int_0^{2\pi} (\psi_t^{2N})^* g_{\mathbb{E}^{2N}}(\partial_\theta, \partial_\theta)^{\frac{1}{2}} d\theta = 2\pi (\psi_t^{2N})^* g_{\mathbb{E}^{2N}}(\partial_\theta, \partial_\theta)^{\frac{1}{2}}. \quad (5.20)$$

We can evaluate this using (5.15) at any θ . So let's take $\theta = 0$ where our eigenfunctions, $\phi_{2j}(0) = 0$ and $\phi_{2j-1}(0) = \pi^{-\frac{1}{2}}$, to see that

$$(\psi_t^{2N})^* g_{\mathbb{E}^{2N}}(\partial_\theta, \partial_\theta) = 2(4\pi)^{\frac{1}{2}} t^{3/2} \sum_{j=1}^{2N} e^{-\lambda_j t} \lambda_j |\phi_j(0)|^2 = 4\pi^{-\frac{1}{2}} t^{3/2} \sum_{j=1}^N e^{-j^2 t} j^2. \quad (5.21)$$

So the lengths of the images of \mathbb{S}^1 are

$$L(\psi_t^{2N}(\mathbb{S}^1)) = 2\pi \left(4\pi^{-\frac{1}{2}} t^{3/2} \sum_{j=1}^N e^{-j^2 t} j^2 \right)^{\frac{1}{2}} \quad (5.22)$$

and the lengths of their limits as $N \rightarrow \infty$ increase to

$$L(\psi_t(\mathbb{S}^1)) = 2\pi \left(4\pi^{-\frac{1}{2}} t^{3/2} \sum_{j=1}^{\infty} e^{-j^2 t} j^2 \right)^{\frac{1}{2}} \quad (5.23)$$

and by (5.13) we have

$$\lim_{t \rightarrow 0^+} L(\psi_t(\mathbb{S}^1)) = 2\pi. \quad (5.24)$$

However if we take $t \rightarrow 0$, before $N \rightarrow \infty$, we get

$$\lim_{t \rightarrow 0^+} L(\psi_t^{2N}(\mathbb{S}^1)) = \lim_{t \rightarrow 0^+} 2\pi \left(4\pi^{-\frac{1}{2}} t^{3/2} \sum_{j=1}^N e^{-j^2 t} j^2 \right)^{\frac{1}{2}} = 0(1^2 + 2^2 + \dots + N^2) = 0.$$

Note that (5.24) can be independently verified using a Theta function as seen in Edwards' text [49]. Consider

$$\Theta(x) = \sum_{n=1}^{\infty} e^{-\pi n^2 x}. \quad (5.25)$$

Then $\Theta(x)$ satisfies the equation

$$\frac{1 + 2\Theta(x)}{1 + 2\Theta(1/x)} = \frac{1}{\sqrt{x}}. \quad (5.26)$$

Differentiating this equation, we obtain the following:

$$-4x^{3/2}\Theta'(x) = 1 + 2\Theta(1/x) + \frac{4}{x}\Theta'(1/x). \quad (5.27)$$

That is,

$$4\pi x^{3/2} \sum_{n=1}^{\infty} n^2 e^{-\pi n^2 x} = 1 + 2\Theta(1/x) + \frac{4}{x}\Theta'(1/x). \quad (5.28)$$

Setting $t = \pi x$, we get

$$4\pi^{-1/2} t^{3/2} \sum_{n=1}^{\infty} n^2 e^{-n^2 t} = 1 + 2\Theta(\pi/t) + \frac{4\pi}{t}\Theta'(\pi/t). \quad (5.29)$$

Note that the right hand side approaches 1 as $t \rightarrow 0$ due to exponential decay of the Theta function. Hence, we have

$$\lim_{t \rightarrow 0^+} L(\psi_t(\mathbb{S}^1)) = 2\pi. \quad (5.30)$$

5.4 Embedding Spheres with a Truncated Heat Kernel

For a round sphere, \mathbb{S}^2 of dimension 2, the first 3 nonconstant eigenfunctions all share the same eigenvalue (as we saw in 3.8). Due to the symmetry, the truncated heat kernel $\psi_t^3(\mathbb{S}^2)$ is an embedding whose image is a round sphere in \mathbb{E}^3 but its radius is smaller than that of the original sphere and decreases to 0 as

$t \rightarrow 0$. In fact, for a round sphere, \mathbb{S}^n of dimension n , the first $n + 1$ nonconstant eigenfunctions all share the same eigenvalue. Thus for $N = n + 1$, the truncated heat kernel, $\psi_t^N(\mathbb{S}^n)$, is an embedding whose image is a round sphere in \mathbb{E}^{n+1} .

If a manifold (M^n, g) is C^2 close to a round sphere \mathbb{S}^n , then its eigenvalues and eigenvectors will be close to that of the sphere. If they are sufficiently close then $\psi_t^N : M^n \rightarrow \mathbb{E}^N$ will be an embedding for $N = n + 1$ as well.

In fact, if (M_j^n, g_j) have Ricci $\geq (n - 1)H$ and $M_j^n \xrightarrow{\text{GH}} \mathbb{S}^n$ then by Cheeger-Colding [25], their eigenfunctions and eigenvalues converge. So perhaps, for j sufficiently large, $\psi_{j,t}^N : M_j \rightarrow \mathbb{E}^N$ is an embedding for $N = n + 1$.

5.5 Embedding Tori with a Truncated Heat Kernel Map

For a flat torus of the form $\mathbb{S}^1 \times \mathbb{S}^1$, we saw in Section 3.6 that the first four nonconstant eigenfunctions are

$$A_{1,0} \sin(\theta), \quad A_{1,0} \cos(\theta), \quad A_{0,1} \cos(\varphi), \quad \text{and} \quad A_{0,1} \sin(\varphi) \tag{5.31}$$

where $A_{1,0} = A_{0,1} = 1/(\pi\sqrt{2})$ with eigenvalue $\lambda_{1,0} = \lambda_{0,1} = 1$ because $A = B = 1$. So the truncated truncated heat kernel map, ψ_t^4 , maps the torus onto

$$\mathbb{S}_R^1 \times \mathbb{S}_R^1 \subset \mathbb{E}^2 \times \mathbb{E}^2 \tag{5.32}$$

where $R = 1/(\pi\sqrt{2})$. Thus the pullback of the Euclidean metric tensor, $\psi_t^{4*} g_{\mathbb{E}^4}$, is just a rescaling of the flat torus metric tensor, $g_{\mathbb{S}^1 \times \mathbb{S}^1}$.

For a flat torus of the form $\mathbb{S}^1 \times \mathbb{S}_{1/10}^1$ we saw in Section 3.7 that the first twenty eigenfunctions only depend upon on the first parameter and so ψ_t^4 even ψ_t^{19} maps into a parametrized circle lying in \mathbb{E}^{19} exactly as in the image of a single \mathbb{S}^1 although it will be scaled differently. The pull back of the metric tensor, $\psi_t^{19*} g_{\mathbb{E}^{19}}$ will not be a positive definite metric tensor on the torus. On a thin torus, we need to take N very large before the pull back becomes positive definite.

5.6 Embedding Tori with Selected Eigenfunctions

Here we embed the thin torus, $\mathbb{S}^1 \times \mathbb{S}_{1/10}^1$, with its metric tensor of the form,

$$g_{\mathbb{S}^1 \times \mathbb{S}_{1/10}^1} = d\theta^2 + (1/10)^2 d\varphi^2, \tag{5.33}$$

into \mathbb{E}^4 using only four eigenfunctions. To do this, we select the four eigenfunctions corresponding to

$$\lambda_{0,1} = \lambda_{10,0} = 100 \tag{5.34}$$

and define the map,

$$\psi_t^{\lambda=100}(\theta, \phi) = \left(\frac{\sqrt{5}}{\pi} \sin(10\theta), \frac{\sqrt{5}}{\pi} \cos(10\theta), \frac{\sqrt{5}}{\pi} \sin(\phi), \frac{\sqrt{5}}{\pi} \cos(\phi) \right). \quad (5.35)$$

We claim that

$$\begin{aligned} (\psi_t^{\lambda=100})^* g_{\mathbb{E}^4} &= (\psi_t^{\lambda=100})^* (dx_1^2 + dx_2^2 + dx_3^2 + dx_4^2) \\ &= \left(10 \frac{\sqrt{5}}{\pi}\right)^2 (\cos^2(10\theta) + \sin^2(10\theta)) d\theta^2 \\ &\quad + \left(\frac{\sqrt{5}}{\pi}\right)^2 (\cos^2(\phi) + \sin^2(\phi)) d\phi^2. \end{aligned}$$

To see this, observe that

$$\begin{aligned} ((\psi_t^{\lambda=100})^* (dx_i))(\partial_\theta) &= dx_i((\psi_t^{\lambda=100})_* \partial_\theta) = ((\psi_t^{\lambda=100})_* \partial_\theta) x_i \\ &= ((\psi_t^{\lambda=100})_* \partial_\theta) x_i = \partial_\theta(x_i \circ \Psi_t^{\lambda=100}). \end{aligned}$$

So we have

$$\begin{aligned} ((\psi_t^{\lambda=100})^* (dx_1))(\partial_\theta) &= \partial_\theta \left(\frac{\sqrt{5}}{\pi} \sin(10\theta) \right) = 10 \frac{\sqrt{5}}{\pi} \cos(10\theta) \\ ((\psi_t^{\lambda=100})^* (dx_2))(\partial_\theta) &= \partial_\theta \left(\frac{\sqrt{5}}{\pi} \cos(10\theta) \right) = -10 \frac{\sqrt{5}}{\pi} \sin(10\theta) \\ ((\psi_t^{\lambda=100})^* (dx_3))(\partial_\phi) &= \partial_\phi \left(\frac{\sqrt{5}}{\pi} \sin(\phi) \right) = \frac{\sqrt{5}}{\pi} \cos(\phi) \\ ((\psi_t^{\lambda=100})^* (dx_4))(\partial_\phi) &= \partial_\phi \left(\frac{\sqrt{5}}{\pi} \cos(\phi) \right) = -\frac{\sqrt{5}}{\pi} \sin(\phi) \end{aligned}$$

and the other pullbacks of dx_i acting on ∂_θ and ∂_ϕ are zero. Thus

$$(\psi^{\lambda=100})^* g_{\mathbb{E}^4} = \left(10 \frac{\sqrt{5}}{\pi}\right)^2 d\theta^2 + \left(\frac{\sqrt{5}}{\pi}\right)^2 d\phi^2 \quad (5.36)$$

is a rescaling of

$$g_{\mathbb{S}^1 \times \mathbb{S}_{1/10}^1} = d\theta^2 + (1/10)^2 d\phi^2. \quad (5.37)$$

This also works to embed $\mathbb{S}^1 \times \mathbb{S}_{1/2}^1$ into \mathbb{E}^4 . We can take the four eigenfunctions corresponding to $\lambda_{0,1} = \lambda_{4,0} = 4$ to define a *selective embedding*:

$$\psi_t^{\lambda=4}(\theta, \phi) = e^{-4t/2} (A_{2,0} \sin(2\theta), A_{2,0} \cos(2\theta), A_{0,1} \sin(\phi), A_{0,1} \cos(\phi)). \quad (5.38)$$

Note that a *selective embedding* as described above is a well chosen projection of the heat kernel embedding in carefully selected directions. This is our own notion introduced vaguely here and is not mentioned in the literature elsewhere.

This will not work for arbitrary $\mathbb{S}_A^1 \times \mathbb{S}_B^1$ with $A > B$ as $\lambda = (1/B)^2$ might only have multiplicity 2 with a pair of eigenfunctions that only depend on φ . In that case we might take a pair of eigenvectors depending on θ with an eigenvalue near $(1/B)^2$ to define an **approximately selective embedding**

$$\psi_t^{\lambda \text{ near } (1/B)^2} : \mathbb{S}_A^1 \times \mathbb{S}_B^1 \rightarrow \mathbb{E}^4. \quad (5.39)$$

Although this would not be a rescaling, it would be biLipschitz close to one.

It would be easier to find a pair of eigenvectors depending on θ with an eigenvalue nearer to $(k/B)^2$ for large k . We might then use these in conjunction with a pair of eigenvectors depending on φ with $\lambda = (k/B)^2$ to define an **approximately selective embedding**

$$\psi_t^{\lambda \text{ near } (k/B)^2} : \mathbb{S}_A^1 \times \mathbb{S}_B^1 \rightarrow \mathbb{E}^4. \quad (5.40)$$

We might even conjecture that that with appropriate rescaling by a well chosen sequence, $\zeta_k \in (0, \infty)$, the pullbacks of these **approximately selective embedding** converge to the standard flat metric on the torus:

$$\lim_{k \rightarrow \infty} \zeta_k (\psi_t^{\lambda \text{ near } (k/B)^2})^* g_{\mathbb{E}^4} = g_{\mathbb{S}_A^1 \times \mathbb{S}_B^1}. \quad (5.41)$$

This would be straightforward to check for tori. Could something like this work for more general classes of manifolds? See Conjecture 8.2 at the end of this paper.

5.7 Estimating Lengths vs Extrinsic Distances

Note that Bérard-Besson-Gallot heat kernel embeddings can be used to approximate the lengths of curves:

$$L_g(C) = \int_0^1 g(C'(s), C'(s))^{1/2} ds. \quad (5.42)$$

If we choose $t < T_\epsilon$ as in (5.17) then

$$|L_{(\psi_t)^* g_{\mathbb{E}^4}}(C) - L_g(C)| < \epsilon L_g(C). \quad (5.43)$$

So

$$(1 - \epsilon) L_g(C) < L_{(\psi_t)^* g_{\mathbb{E}^4}}(C) < (1 + \epsilon) L_g(C). \quad (5.44)$$

The Riemannian distance between points $x, y \in M$ is found by taking

$$d_M(x, y) = \inf\{L(C) : C(0) = x, C(1) = y, C : [0, 1] \rightarrow M\}. \quad (5.45)$$

So

$$(1 - \epsilon) d_M(x, y) < d_{\psi_t(M)}(\psi_t(x), \psi_t(y)) < (1 + \epsilon) d_M(x, y). \quad (5.46)$$

So ψ_t is $(1 + \epsilon)$ -biLipschitz with respect to the intrinsic metric of the image $\psi_t(M)$ but not with respect to the metric d_{ℓ^2} .

To make this more clear, let's use the truncated heat kernel embeddings. Taking $N > N_{t,\epsilon}$ so that we have (5.18), we can repeat the above to see that

$$(1 - \epsilon - \epsilon)L_g(C) < L_{(\psi_t^N)^* g_{\mathbb{R}^N}}(C) < (1 + \epsilon + \epsilon)L_g(C) \quad (5.47)$$

and

$$(1 - \epsilon - \epsilon) d_M(x, y) < d_{\psi_t^N(M)}(\psi_t(x), \psi_t(y)) < (1 + \epsilon + \epsilon) d_M(x, y). \quad (5.48)$$

So ψ_t is $(1 + \epsilon + \epsilon)$ -biLipschitz with respect to the intrinsic distance within the image $\psi_t^N(M)$ but not with respect to the extrinsic Euclidean distance.

In fact, the extrinsic Euclidean distance between the points in the image are converging to 0. This can be seen as follows:

$$d_{\mathbb{R}^N}(\psi_t(x), \psi_t(y)) \leq d_{\mathbb{R}^N}(\psi_t(x), 0) + d_{\mathbb{R}^N}(0, \psi_t(y)) \quad (5.49)$$

and

$$\begin{aligned} d_{\mathbb{R}^N}(\psi_t(x), 0)^2 &= \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 \sum_{j=1}^N e^{-\lambda_j t} (\phi_j(x))^2 \\ &= \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 h_t^N(x, x) \end{aligned}$$

Taking $N \rightarrow \infty$ first we get

$$d_{\ell^2}(\psi_t(x), \psi_t(x))^2 = \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 h_t(x, x). \quad (5.50)$$

By the Minakshisundaram-Pleijel expansion in (5.10)-(5.11), we can see that,

$$\lim_{t \rightarrow 0} \left(\sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \right)^2 h_t(x, x) = \lim_{t \rightarrow 0} 2(4\pi)^{n/2} t^{n/2+1} \frac{e^{-d_M^2(x,x)/(4t)}}{(4\pi t)^{n/2}} = 0. \quad (5.51)$$

Thus

$$\lim_{t \rightarrow 0} \lim_{N \rightarrow \infty} d_{\mathbb{R}^N}(\psi_t^N(x), \psi_t^N(y)) = \lim_{t \rightarrow 0} d_{\ell^2}(\psi_t(x), \psi_t(y)) = 0. \quad (5.52)$$

5.8 Other Rescalings of the Heat Kernel Embedding

Bérard-Besson-Gallot defined an embedding where they rescaled by volume

$$I_t(x) = \sqrt{\text{Vol}(M)}\Psi_t(x), \quad (5.53)$$

in Part III of [16].

They also defined an embedding into an infinite dimensional sphere in Part VIII of [16].

$$K_t(x) = \Psi_t(x)/|\Psi_t(x)|. \quad (5.54)$$

They proved the pull back satisfies

$$K_t^* g_{\ell^2} = 1/(2t) \left(g - (t/3)\text{Ricci} + O(t^2) \right) \text{ as } t \rightarrow 0^+. \quad (5.55)$$

5.9 Spectral Convergence of Riemannian Manifolds

Inspired by the work of Gromov [55] and of Fukaya [53], Bérard-Besson-Gallot defined a spectral distance between Riemannian manifolds.

To define a *spectral distance* between two Riemannian manifolds, M_1 and M_2 , they took the Hausdorff distance between $I_t(M_1)$ and $I_t(M_2)$ of (5.53) and considered a minmax over various reorderings of the eigenfunctions in Part IV of [16]. They proved a compactness theorem for their spectral distance and they prove that when sequences of manifolds converge with respect to this spectral distance their eigenvalues converge.

Kasue and Kumura defined a different spectral convergence in [61] [62] [60]. Their notion is related to this one but is defined in a completely different way using maps that almost preserve heat kernels rather than Hausdorff distances between embeddings. They also proved that when sequences of manifolds converge with respect to this spectral distance their eigenvalues converge.

6 Applications to Dimension Reduction of High Dimensional Data Sets

Heat kernel embeddings have been applied to achieve *dimension reduction*: the mapping of data sets of points in high dimensional spaces to lower dimensional spaces. An example of such a data set would be a collection of photographs. Each photograph would be a point. See for example Figure 6.1. If the photographs are in grey scale with $1,000 \times 1,000$ pixels, then their points lie in $X = \mathbb{R}^{1,000,000}$. In general, the data set is described by $\{x_1, \dots, x_k\} \in X$, where X is

the high dimensional space endowed with a distance, d_X , that is meaningful to the application.

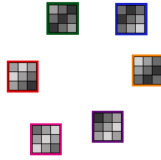


Figure 6.1: Here we have 6 photos each of which has 9 pixels, so they may be viewed as points in \mathbb{R}^9 . For example, the leftmost photo is $(2, 1, 2, 1, 2, 3, 2, 3, 4)$ and the one above it is $(2, 3, 4, 3, 4, 3, 4, 3, 2)$. The distance between these photos measured using the Euclidean distance is then $\sqrt{24}$.

The distances between the points can be stored as a **distance matrix**, $d_X(i, j)$. Sometimes the distances are just the extrinsic Euclidean distances. A neighborhood graph can be created by joining some nearby points with edges. One method is to select the k nearest neighbors (kNN) for each point to be connected with an edge so the resulting graph has degree k : k edges per vertex. Another method is to connect each point with every point in an ϵ neighborhood about it. See Figure 6.2. A distance of shortest paths can then be found using an algorithm by Dijkstra [44] or we can use the extrinsic Euclidean distance.

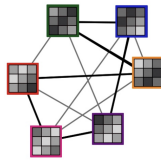


Figure 6.2: We have created a graph with a photo at each vertex. Here we have connected photos that are within an $\epsilon = 5$, so that there are edges between pairs of photos that are a distance less than 5 apart. Note it turns out there are 4 edges per vertex, so this is also a kNN graph with $k = 4$. Note that the thick lines connect vertices that are a distance 3 apart and the thinner lines are for the vertices that are further apart.

In 2000, Tenenbaum-deSilva-Langford [91] created the ISOMap algorithm combining these ideas with Multidimensional Scaling (MDS) which is a technique from Linear Algebra. Roweis-Saul [82] introduced Nonlinear Dimension-

ality Reduction using the Locally Linear Embedding (LLE) algorithm which eliminated the need to estimate pairwise distances between widely separated points. These initial papers introduced algorithms and demonstrated their effectiveness in studying sample data sets but did not include proofs.

When we assume the data set of points lies on a submanifold of the high dimensional space, the process of dimension reduction is called *manifold learning*. For example, we might have a collection of photos from a continuous film, in which case they lie on a curve which is a 1 dimensional manifold. Or they might be photos of the same scene taken from many directions at the same distances, in which case the underlying submanifold would be $SO(3)$. Note that the submanifold is not known: it is just presumed to exist, and then if the right technique is used, its properties allow us to reduce the dimension.

The goal is to map the data points into a lower dimensional space and thus discover the underlying structure. Ideally we might even preserve the distances under this map (just as in a heat kernel embedding), although even just obtaining an embedding is of interest. The heat kernel is also useful because it is robust to noise. In 2001, Belkin and Niyogi introduced the idea of embedding data sets using *eigenmaps* whose components are eigenfunctions [13]. In this section we describe their approach and then later we describe the *diffusion maps* of Coifman-Lafon whose components are weighted eigenfunctions [38]. See Figure 6.3.

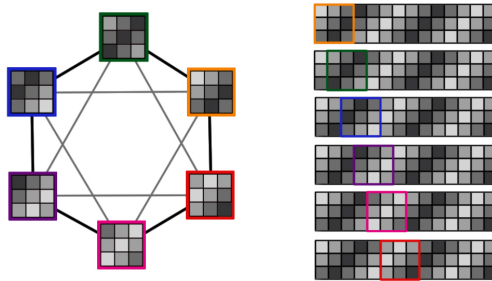


Figure 6.3: On the left is the image of the data set from Figure 6.1 after it has been mapped into \mathbb{R}^2 by an eigenmap. We have placed each photo on top of its image under the eigenmap of Coifman-Lafon and we then drew the edges connecting them. We see that the data is now arranged symmetrically into a circle. On the right, we have used the circular symmetry to glue the photos together to form a panoramic image which is periodic. This symmetry was not visible in the original disorganized data set.

6.1 Belkin-Niyogi Eigenmaps of Data Sets

In 2001, Belkin and Niyogi introduced *Laplacian eigenmaps* [13] as a method to provide dimension reduction for data sets that lie on submanifolds in a high dimensional space, X . Their algorithm has three steps as follows:

Algorithm:

Step 1: Given k points, $\{x_1, \dots, x_k\} \subset X$, they constructed an adjacency graph connecting neighboring points, $x_i \sim x_j$, either using an ϵ -neighborhood approach or the k nearest neighbors (kNN).

Step 2: They created an $k \times k$ **heat kernel matrix**,

$$W_{i,j}^t = e^{-|x_i - x_j|^2/t} \text{ for } x_i \sim x_j \text{ and } W_{i,j}^t = 0 \text{ otherwise.} \quad (6.1)$$

In particular, $W_{i,i}^t = e^{-0^2/t} = 1$ because $x_i \sim x_i$. They define a **diagonal degree matrix**

$$D_{i,i} = \sum_{j=1}^n W_{i,j} \text{ where } W_{i,j} = W_{i,j}^t, \quad (6.2)$$

a **graph Laplacian matrix**, $L = D - W$.

Step 3: Belkin-Niyogi define an **eigenmap**: $V_t^N : \{x_1, \dots, x_k\} \rightarrow \mathbb{R}^N$, for any $N \leq k - 1$ by

$$V_t^N(x_i) = (v_1(i), v_2(i), \dots, v_N(i)) \in \mathbb{R}^N \text{ where } v_j(i) \text{ is the } i^{\text{th}} \text{ component of } v_j. \quad (6.3)$$

where $0 = \mu_0 < \mu_1 \leq \dots \leq \mu_{k-1}$ are the eigenvalues and

$$v_0, v_1, \dots, v_{k-1} \in \mathbb{R}^k \quad (6.4)$$

is an orthonormal collection of **eigenvectors** of the **normalized graph Laplacian** $D^{-1}L$:

$$D^{-1}L v_j = \mu_j v_j. \quad (6.5)$$

Belkin and Niyogi leave out the 0^{th} eigenvector v_0 in the definition of their eigenmap V_t^N because all of its entries are the same. This can easily be seen by observing that the sum of the entries in each row of $D^{-1}L$ is 0:

$$\sum_{j=1}^k (D^{-1}L)_{i,j} = \sum_{j=1}^k D_i^{-1} (D - W)_{i,j} = \sum_{j=1}^k I_{i,j} - D_i^{-1} \sum_{j=1}^k W_{i,j} = 1 - D_i^{-1} D_i = 0. \quad (6.6)$$

In [13], Belkin and Niyogi explain that intuitively if $\{x_1, \dots, x_k\} \subset X$ lie as a uniformly distributed collection of points in a submanifold, M^n , of small dimension, n , in the high dimensional Euclidean space, X , then they can use eigenfunctions, $\phi_j : M \rightarrow \mathbb{R}$, of the Laplacian, Δ , on M to define a **smooth eigenmap**,

$$\Phi^N : M^n \rightarrow \mathbb{R}^N \text{ as } \Phi^N(x_i) = (\phi_1(x_i), \phi_2(x_i), \dots, \phi_N(x_i)). \quad (6.7)$$

Note that this **smooth eigenmap**, $\Phi^N(x)$, is a truncation of an unweighted version of Bérard-Besson-Gallot's heat kernel embedding, $\Psi_t(x)$, that we saw in (5.4) [16].

Belkin and Niyogi's three step algorithm defines an eigenmap, V_t^N , on the data set, $\{x_1, \dots, x_k\}$, without knowing the ambient manifold, M , such that

$$\Phi^N(x_i) = (\psi_1(x_i), \psi_2(x_i), \dots, \psi_N(x_i)) \approx (v_1(i), v_2(i), \dots, v_N(i)) = V_t^N(x_i) \quad (6.8)$$

The approximation $\Phi^N(x_i) \approx V_t^N(x_i)$ was not proven rigorously in Belkin and Niyogi's early papers however they did provide some of intuition leading to their algorithm.

The intuitive reasoning behind Belkin and Niyogi's choice of matrix in (6.1) was explained by Belkin roughly as follows in his doctoral dissertation [12]. Recall that given a function, $f : M \rightarrow \mathbb{R}$, then

$$u(x, t) = \int_M h_t(x, y) f(y) \, dvol_y \quad (6.9)$$

satisfies

$$\partial_t u = \Delta u \text{ and } \lim_{t \rightarrow 0^+} u(x, t) = f(x). \quad (6.10)$$

Thus

$$\begin{aligned} (\Delta f)(x) &= (\Delta u)(x, 0) = (\partial_t u)(x, 0) = \lim_{t \rightarrow 0^+} \frac{-1}{t} (u(x, 0) - u(x, t)) \\ &= \lim_{t \rightarrow 0^+} \frac{-1}{t} \left(f(x) - \int_M h_t(x, y) f(y) \, dvol_y \right) \end{aligned}$$

Since they are studying a data set, Belkin and Niyogi ignored the limit as $t \rightarrow 0^+$ and consider integration to be summation:

$$(\Delta f)(x_i) \approx f(x_i) - \sum_{j=1}^k h_t(x_i, x_j) f(x_j) \quad (6.11)$$

They decided to choose a weight matrix $W_{i,j}$ so that

$$D_{i,i}^{-1} W_{i,j} \approx h_t(x_i, x_j) \quad (6.12)$$

so that

$$(\Delta f)(x_i) \approx f(x_i) - \sum_{j=1}^k (D^{-1}W)_{i,j} f(x_j) = (I - D^{-1}W)_{i,j} f(x_j). \tag{6.13}$$

In this way, the eigenfunctions $\phi_j(x)$ of the Laplace Beltrami operator, $\Delta\phi_j = -\lambda_j\phi_j$, seem to correspond with the eigenvectors, v_j , of $I - D^{-1}W = D^{-1}L$ as in (6.5) so that

$$\phi_j(x_i) \approx v_j(i) \quad \forall i, j \in \{1, \dots, k\} \tag{6.14}$$

which intuitively justifies (6.8). Note that they made no claim that the eigenvalues match and in practice the eigenvalues, λ_j and μ_j are not close even as $t \rightarrow 0^+$ as we will see in upcoming sections.

Note that by (6.12) and the definition of the degree matrix (6.2), they achieve the normalization of the heat kernel:

$$1 = \int_M h_t(x, y) dy \approx \sum_{j=1}^k h_t(x_i, x_j) = \sum_{j=1}^k D_i^{-1} W_{i,j} = 1. \tag{6.15}$$

Finally Belkin-Niyogi approximated the heat kernel as we've seen throughout by

$$h_t(x_i, x_j) \approx \frac{1}{(4\pi t)^{n/2}} e^{-|x_i - y_i|^2 / (4t)} \tag{6.16}$$

to justify their choice of $W_{i,j}$ in (6.1) although their choice has division by t instead of $4t$ in the exponent. Since this just rescales time, it does not cause any issues for them.

A more rigorous discussion appears in the 2006 papers of Belkin-Niyogi using a probabilistic approach [11] [14]. See also Kondor-Lafferty work in [63]. See also their joint work with Sindhwani in [15].

6.2 A Brief History of Graph Laplacians starting with Dodziuk

It should be noted that the graph Laplacian was first defined by Dodziuk in 1984 [47]. His graph Laplacian was on a simple graph without weighted edges. He used the **difference Laplacian**,

$$“\Delta” f(x_i) = \sum_{x_j \sim x_i} (f(x_j) - f(x_i)) = -D_i f(x_i) + \sum_{x_j \sim x_i} f(x_j) \tag{6.17}$$

that had been studied on square lattices in Euclidean space as early as 1928 by Courant, Friedrichs and Lewy [40]. Here D_i is the degree of vertex x_i and is equal

to the number of vertices x_j such that $x_j \sim x_i$. Defining W to be the adjacency matrix such that $W_{i,j} = 1$ when $x_i = x_j$ or there is an edge between x_i and x_j , and 0 otherwise, we see that Dodziuk's graph Laplacian satisfies,

$$\begin{aligned}
 \text{"}\Delta\text{"} f(x_i) &= \sum_{x_j \sim x_i} (f(x_j) - f(x_i)) = \sum_{j=1}^k (W - D)_{i,j} f(x_j) \quad (6.18)
 \end{aligned}$$

where D is the diagonal degree matrix $D_{i,i} = D_i$. Note the opposite sign on this version. In his paper, Dodziuk proved the maximum principle, the Harnack inequality, and Cheeger's bound for the lowest eigenvalue all extend to the graph setting in [47]. In 1988, he developed a Green's formula with Karp in [48]. For early applications of these ideas to Markov chains see the paper of Lawler and Sokol [65].

6.3 Continuing the History of Graph Laplacians with work of Chung

A deep analysis of graph Laplacians appears in the work of Chung in [34] where she studies weighted graphs with a symmetric weight matrix, W . The degree of the vertex is then defined to be $D_i = \sum_{j=1}^k W_{i,j}$. Chung's Graph Laplacian matrix is defined $L = D - W$ with the opposite sign used by Dodziuk. She defines a normalized graph Laplacian matrix, $D^{-1}L$, which might not be symmetric. So she defines a normalized symmetrized graph Laplacian matrix,

$$\mathcal{L} = D^{1/2} (D^{-1}L) D^{-1/2} = D^{-1/2} L D^{-1/2} = I - D^{-1/2} W D^{-1/2}, \quad (6.19)$$

which is a symmetric matrix. Taking a function $f : \{x_1, \dots, x_k\} \rightarrow \mathbb{R}$, she defines $(\mathcal{L}f) : \{x_1, \dots, x_k\} \rightarrow \mathbb{R}$ as follows:

$$(\mathcal{L}f)(x_i) = \sum_{j=1}^k (I - D^{-1/2} W D^{-1/2})_{i,j} f(x_j), \quad (6.20)$$

so that

$$(\mathcal{L}f)(x_i) = f(x_i) - \sum_{j=1}^k D_i^{-1/2} W_{i,j} D_j^{-1/2} f(x_j). \quad (6.21)$$

Note that the eigenvalues μ_j of the normalized graph Laplacian, $D^{-1}L$, are also the eigenvalues of \mathcal{L} because

$$\begin{aligned}
 \mathcal{L}(D^{1/2}v_j) &= D^{-1/2} L D^{-1/2} (D^{1/2}v_j) = D^{1/2} D^{-1} L v_j \\
 &= D^{1/2} (-\mu_j v_j) = -\mu_j (D^{1/2}v_j).
 \end{aligned}$$

Because of this well known relationship, some people redefine Belkin-Niyogi's eigenmaps using the eigenvectors of the symmetrized normalized Laplacian \mathcal{L} .

In 1995-1997 Chung completed joint work with Yau on Graph eigenvalues and Sobolev inequalities [35], joint work with Grigor'yan and Yau proving upper bounds on eigenvalues of finite graphs [32], and joint work with Yau on the heat kernel and its trace for graphs [36]. She has applied these ideas in many directions including developing a pagerank for web search engines [37].

Note that Belkin-Niyogi's graph Laplacian is defined using their own weight matrix W that was not studied by Chung. It does not agree with the adjacency matrix of Dodziuk even when all edges have the same weight. Their graph Laplacian only fits with Chung's work in the sense that it uses her notion of a weighted graph Laplacian (as in 1.4 of her textbook [31]). It is not the heat kernel matrix defined in Chung's work with Yau in [36] and explored further in the later chapters of her textbook [31].

6.4 Graph Laplacians Approximating Smooth Laplacians

Hein, Audibert, and Luxburg provide a rigorous discussion of the relationship between graph Laplacians to manifold Laplacians in [56]. Singer improves the convergence rate from graph to manifold Laplacians in [84] by also assuming uniformly distributed points.

In [95], Wardetzky, Mathur, Kalberer, and Grinspun study five different discrete Laplacian and discuss which properties each shares with the continuous Laplacian. They conclude that no discrete Laplacian can have all the properties but that different ones may be used to achieve different properties.

Most recently, Burago-Ivanov-Kurylev have shown in [19] that eigenvalues and eigenfunctions of the Laplace-Beltrami operator on a Riemannian manifold are approximated by eigenvalues and eigenvectors of a (suitably weighted) graph Laplace operator of a proximity graph on an epsilon-net. Earlier work by Fujiwara in [52] had also considered ϵ -nets but he only approximated the eigenvalues and his estimates were not as strong. There is also unpublished work of Aubry from 2014 providing estimates that depend only on the upper bounds on diameter and sectional curvature of M and lower bounds on the injectivity radius.

6.5 An Eigenmap of a Hexagon

As a simple case let us work out an eigenmap of 6 points, $\{x_1, x_2, x_3, x_4, x_5, x_6\}$, where

$$x_j = (\cos(2\pi j/6), \sin(2\pi j/6)) \quad (6.22)$$

lie on the unit circle. We can consider kNN with $k = 2$ nearest points or ϵ slightly > 1 to determine the same collection of unit length edges, so that our graph is the standard hexagon with two edges of length one attached at every vertex. Then by (6.1), we have

$$W = \begin{pmatrix} 1 & e^{-1/t} & 0 & 0 & 0 & e^{-1/t} \\ e^{-1/t} & 1 & e^{-1/t} & 0 & 0 & 0 \\ 0 & e^{-1/t} & 1 & e^{-1/t} & 0 & 0 \\ 0 & 0 & e^{-1/t} & 1 & e^{-1/t} & 0 \\ 0 & 0 & 0 & e^{-1/t} & 1 & e^{-1/t} \\ e^{-1/t} & 0 & 0 & 0 & e^{-1/t} & 1 \end{pmatrix}. \quad (6.23)$$

and our diagonal matrix D has $D_{i,i} = \sum_{j=1}^n W_{i,j} = 1 + 2e^{-1/t}$. So

$$L = D - W = \begin{pmatrix} 2e^{-1/t} & -e^{-1/t} & 0 & 0 & 0 & -e^{-1/t} \\ -e^{-1/t} & 2e^{-1/t} & -e^{-1/t} & 0 & 0 & 0 \\ 0 & -e^{-1/t} & 2e^{-1/t} & -e^{-1/t} & 0 & 0 \\ 0 & 0 & -e^{-1/t} & 2e^{-1/t} & -e^{-1/t} & 0 \\ 0 & 0 & 0 & -e^{-1/t} & 2e^{-1/t} & -e^{-1/t} \\ -e^{-1/t} & 0 & 0 & 0 & -e^{-1/t} & 2e^{-1/t} \end{pmatrix}. \quad (6.24)$$

Here D is a multiple of the identity, so we have $\mathcal{L} = D^{-1/2} L D^{-1/2} = D^{-1} L$, and

$$\mathcal{L} = D^{-1} L = \frac{2e^{-1/t}}{1 + 2e^{-1/t}} \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 1 \end{pmatrix} \quad (6.25)$$

Thus we will get the same eigenmaps, $V_t^N : \{x_1, \dots, x_k\} \rightarrow \mathbb{R}^N$, defined using Belkin-Niyogi's $D^{-1}L$ as we get using the symmetrized \mathcal{L} to compute the eigenvectors.

We know the $\lambda_0 = 0$ eigenfunction is constant and quickly confirm the same for eigenvectors of \mathcal{L} :

$$\frac{2e^{-1/t}}{1 + 2e^{-1/t}} \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = 0 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad (6.26)$$

Ordinarily we would next compute the eigenvectors numerically, but here we can quickly guess them using Belkin-Niyogi's intuition as described in (6.8): $v_j(i) \approx \phi_j(x_i)$.

In Section 5.3, we found the eigenfunctions

$$\phi_1(\theta_i) = A_1 \cos(\theta_i) \text{ and } \phi_2(\theta_i) = A_1 \sin(\theta_i). \quad (6.27)$$

By (6.8), we can guess that the next two eigenvectors, v_1 and v_2 , are

$$\begin{pmatrix} v_1(1) \\ v_1(2) \\ v_1(3) \\ v_1(4) \\ v_1(5) \\ v_1(6) \end{pmatrix} = \begin{pmatrix} \cos(2\pi/6) \\ \cos(2\pi 2/6) \\ \cos(2\pi 3/6) \\ \cos(2\pi 4/6) \\ \cos(2\pi 5/6) \\ \cos(2\pi 6/6) \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} \text{ and } \begin{pmatrix} v_2(1) \\ v_2(2) \\ v_2(3) \\ v_2(4) \\ v_2(5) \\ v_2(6) \end{pmatrix} = \begin{pmatrix} \sin(2\pi/6) \\ \sin(2\pi 2/6) \\ \sin(2\pi 3/6) \\ \sin(2\pi 4/6) \\ \sin(2\pi 5/6) \\ \sin(2\pi 6/6) \end{pmatrix} = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} \\ 0 \\ -\frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} \\ 0 \end{pmatrix}$$

where $\sqrt{3}$ is needed to make them orthonormal. Checking the first we see:

$$\begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & 0 & 0 & 0 & -\frac{1}{2} & 1 \end{pmatrix} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} + \frac{1}{4} - \frac{1}{2} \\ -\frac{1}{4} - \frac{1}{2} + \frac{1}{2} \\ \frac{1}{4} - 1 + \frac{1}{4} \\ \frac{1}{2} - \frac{1}{2} - \frac{1}{4} \\ \frac{1}{4} + \frac{1}{2} - \frac{1}{2} \\ -\frac{1}{4} - \frac{1}{4} + 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{4} \\ -\frac{1}{4} \\ -\frac{1}{2} \\ -\frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{2} \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ -1 \\ -\frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

so that

$$\mathcal{L} v_1 = \frac{2e^{-1/t}}{1 + 2e^{-1/t}} \frac{1}{2} v_1. \quad (6.28)$$

Note that the eigenvalue itself does not match the eigenvalue of the Laplacian even though the eigenvector matches the eigenfunction so well.

Next we see that the truncated eigenmap, with $N = 2$, is just a rescaling of the heat kernel embedding, with $N = 2$:

$$V_t^2(x_i) = (v_1(i), v_2(i)) = (\cos(2\pi i/6), \sin(2\pi i/6)) = e^t \Psi_t^2(x_i). \quad (6.29)$$

The fact that we have equality instead of an approximate equality is due to the symmetry of the circle and the points lying on the circle.

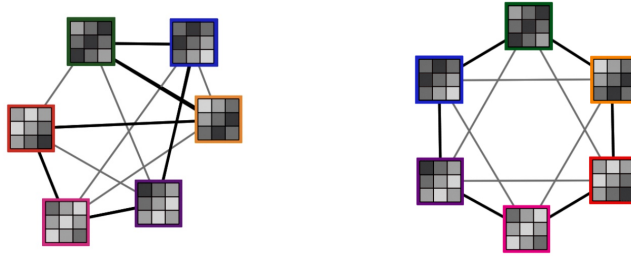


Figure 6.4: On the left is the data set in \mathbb{R}^9 with the points marked by their 3×3 photos and on the right is the image under the truncated eigenmap, V_t^2 , with the image points marked by their 3×3 photos lying on $\mathbb{S}^1 \subset \mathbb{R}^2$.

6.6 Finding the Eigenmap of a Data Set of 3×3 Photos

At the beginning of this section we described the process of dimension reduction using a collection of six 3×3 pixel photos. See Figure 6.4. We will now describe exactly how this is done using a Belkin-Niyogi eigenmap.

We label the original collection of six 3×3 pixel photos on the left side of Figure 6.4 starting with the rightmost and proceeding counter clockwise:

$$\begin{array}{c}
 \begin{array}{l} (1) \\ 2 \\ 3 \\ 2 \\ 3 \\ 4 \\ 3 \\ 4 \\ (3) \end{array} \\
 p_1 =
 \end{array}
 \begin{array}{c}
 \begin{array}{l} (3) \\ 4 \\ 3 \\ 4 \\ 3 \\ 2 \\ 3 \\ 2 \\ (1) \end{array} \\
 p_2 =
 \end{array}
 \begin{array}{c}
 \begin{array}{l} (2) \\ 3 \\ 4 \\ 3 \\ 3 \\ 4 \\ 3 \\ 3 \\ (2) \end{array} \\
 p_3 =
 \end{array}
 \begin{array}{c}
 \begin{array}{l} (2) \\ 1 \\ 2 \\ 1 \\ 2 \\ 3 \\ 2 \\ 3 \\ (4) \end{array} \\
 p_4 =
 \end{array}
 \begin{array}{c}
 \begin{array}{l} (3) \\ 2 \\ 1 \\ 2 \\ 2 \\ 1 \\ 2 \\ 2 \\ (3) \end{array} \\
 p_5 =
 \end{array}
 \begin{array}{c}
 \begin{array}{l} (4) \\ 3 \\ 2 \\ 3 \\ 2 \\ 1 \\ 2 \\ 1 \\ (2) \end{array} \\
 p_6 =
 \end{array}$$

Evaluating the distances between these points we see that some have distance $\sqrt{24}$ and others have every pixel off by 1 so their distance is $\sqrt{9} = 3$. Taking $t = 1$ and observing that $e^{-24} \approx 0$ and $e^{-3} \approx .05$ we get a 6×6 matrix

$$W = \begin{pmatrix} 1 & 0 & .05 & .05 & 0 & 0 \\ 0 & 1 & .05 & 0 & 0 & .05 \\ .05 & .05 & 1 & 0 & 0 & 0 \\ .05 & 0 & 0 & 1 & .05 & 0 \\ 0 & 0 & 0 & .05 & 1 & .05 \\ 0 & .05 & 0 & 0 & .05 & 1 \end{pmatrix} \tag{6.30}$$

Thus our diagonal matrix has $D_{i,i} = 1 + .05 + .05 = 1.1 = 11/10$ so

$$L = D - W = \begin{pmatrix} .1 & 0 & -.05 & -.05 & 0 & 0 \\ 0 & .1 & -.05 & 0 & 0 & -.05 \\ -.05 & -.05 & .1 & 0 & 0 & 0 \\ -.05 & 0 & 0 & .1 & -.05 & 0 \\ 0 & 0 & 0 & -.05 & .1 & -.05 \\ 0 & -.05 & 0 & 0 & -.05 & .1 \end{pmatrix} \quad (6.31)$$

and since $D_{i,i}^{-\frac{1}{2}} = (11/10)^{-1/2} = \sqrt{10/11}$ we have

$$\mathcal{L} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}} = \frac{10}{11} \begin{pmatrix} 1 & 0 & -1/2 & -1/2 & 0 & 0 \\ 0 & 1 & -1/2 & 0 & 0 & -1/2 \\ -1/2 & -1/2 & 1 & 0 & 0 & 0 \\ -1/2 & 0 & 0 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & -1/2 & 1 & -1/2 \\ 0 & -1/2 & 0 & 0 & -1/2 & 1 \end{pmatrix} \quad (6.32)$$

It is easy to check that

$$v_1 = \begin{pmatrix} \frac{\sqrt{3}}{2} \\ -\frac{\sqrt{3}}{2} \\ 0 \\ \frac{\sqrt{3}}{2} \\ 0 \\ -\frac{\sqrt{3}}{2} \end{pmatrix} \text{ and } v_2 = \begin{pmatrix} 1/2 \\ 1/2 \\ 1 \\ -1/2 \\ -1 \\ -1/2 \end{pmatrix} \quad (6.33)$$

are perpendicular eigenfunctions with eigenvalue $= -10/22$. Thus up to scaling, Belkin-Niyogi's eigenmap, $V_t^2(p_i) = (v_1(p_i), v_2(p_i))$, takes the following values:

$$\begin{aligned} V_2(p_1) &= (\sqrt{3}/2, 1/2) & V_2(p_2) &= (-\sqrt{3}/2, 1/2) & V_2(p_3) &= (0, 1) \\ V_2(p_4) &= (\sqrt{3}/2, -1/2) & V_2(p_5) &= (0, -1) & V_2(p_6) &= (-\sqrt{3}/2, -1/2) \end{aligned}$$

which reveals the circular structure of the image seen on the right in Figure 6.4.

6.7 An Eigenmap of a Sphere

Our 2020 student research team (Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara) tried out the Belkin-Niyogi method of eigenmaps using

MATLAB. Recall that in Section 3.8 we observed that the first three nonconstant eigenfunctions of \mathbb{S}^2 up to scaling are

$$\phi_1(x) = X(x) \quad \phi_2(x) = Y(x) \quad \phi_3(x) = Z(x) \quad (6.34)$$

so the image of the heat kernel embedding, $\Psi_t^3 : \mathbb{S}^2 \rightarrow \mathbb{E}^3$, is a round sphere. So this was a good example to test their code on an evenly distributed data set of points in \mathbb{S}^2 to produce an eigenmap V_t^3 to \mathbb{E}^3 .

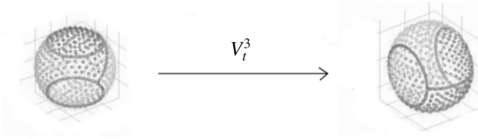


Figure 6.5: Here we see a data set of evenly distributed points on a sphere mapped to a sphere by a truncated eigenmap, V_t^3 , and we see that it agrees up to scaling with the image under a heat kernel embedding, $\psi_t^3 : \mathbb{S}^2 \rightarrow \mathbb{E}^3$. MATLAB computation by Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara.

To get the image to look like a sphere, the points had to be evenly distributed. Ultimately the team devised a collection of biLipschitz disjoint charts to cover a sphere, used those charts to map a triangular lattice from a plane into the sphere, and then added extra points along the edges of the charts, to distribute the points as evenly as possible.

In Figure 6.5 we see the evenly distributed points on the sphere as our data set on the left, and then we see the image under V_t^3 on the right. Since $N = 3$ this agrees up to scaling with the image under the truncation of the sphere's heat kernel embedding, ψ_t^3 , that we studied in Section 3.8. It has been rotated because the eigenvectors were selected in a different order by MATLAB so that up to scaling,

$$V_t^3(x_j) = (v_1(i), v_2(i), v_3(i)) \quad \leftrightarrow \quad \psi_t^3(x_j) = (-Z(x_j), -X(x_j), Y(x_j)). \quad (6.35)$$

This captures the concern Bérard-Besson-Gallot mentioned regarding choice of eigenfunctions leading to rotations of the image of the heat kernel embeddings by elements of $\text{SO}(N)$ that we discussed at the end of Section 5.2.

6.8 Eigenmaps of Barbells

Our 2020 student research team (Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara) next applied the the method of eigenmaps to study what the images of heat kernel embeddings of deformed spheres shaped like barbells might look like. In Figure 6.6, there are evenly distributed points on various barbells: *pairs of spheres of radius, 20, with caps removed and cylinders of radius, 5, of various lengths, $L = 100$, $L = 10$ and $L = 1$, glued between them.*

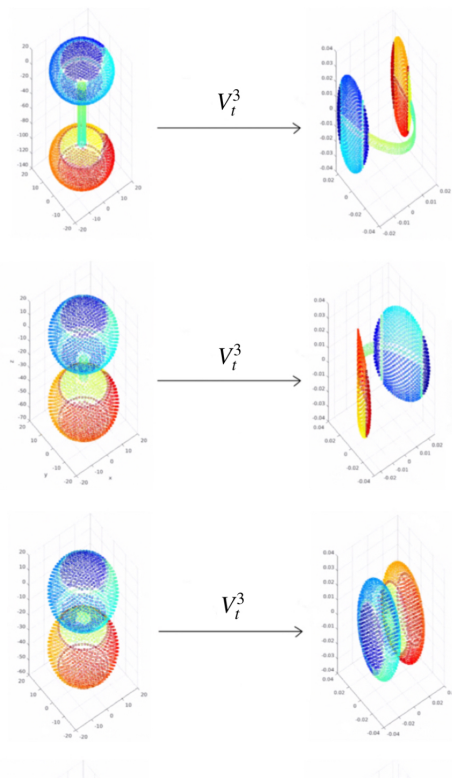


Figure 6.6: Here we see a data set of evenly distributed points on various barbells mapped by an eigenmap, V_t^3 , which gives some possible insight into the image under a heat kernel embedding, Ψ_t^3 of these barbells. MATLAB computation by Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara.

When the cylinder is long, $L = 100$ as on the top of Figure 6.6, there is a first nonzero eigenvalue corresponding to the full height of the manifold which has an eigenfunction that looks something like $\cos(z)$. The next two eigenfunctions

seem to correspond to Z and X . That is the points x_i are mapped to $V_t^3(x_i) = (\cos(Z(x_i)), Z(x_i), X(x_i))$ and so the image of ψ_t^3 is probably not an embedding. It is flattened and then bent around a cosine curve. The students called this image a *headset*.

In the second row of Figure 6.6, the cylinder only has $L = 10$ and yet the same phenomenon has occurred. The image is once again not an embedding and takes the form of a headset. The headset is rotated the points x_i seem to be mapped to

$$V_t^3(x_i) = (-\cos(Z(x_i)), -Z(x_i), X(x_i)) \quad (6.36)$$

up to scaling because different eigenvectors were selected first by MATLAB.

In the bottom row of Figure 6.6, the cylinder is finally short enough that the eigenfunctions correspond to something like X , Y and Z , so the image is not flattened, and the eigenmap is an embedding.

Please note that these are images under the Belkin-Noyogi eigenmaps, V_t^3 , not the images under the smooth eigenmaps, Φ_t^3 , nor the heat embeddings, Ψ_t^3 of Bérard-Besson-Gallot. Nevertheless there is an intuitive relationship between these concepts.

It should be noted that, to save on computation time, the students did not construct an adjacency graph but used the Euclidean distances to define their matrix W (since points that were not close lead to very small entries in W anyway). Note also that shortest paths cutting across lattices of points are not approximating the lengths of direct line segments, so it is not clear that these are better than Riemannian lengths when trying to approximate the heat kernel of a Laplacian.

6.9 Eigenmaps of Lollipops

Our 2020 student research team (Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara) next applied the the method of eigenmaps to study what the heat kernel embeddings of deformed spheres shaped like lollipops might look like. In Figure 6.7, they have evenly distributed points on various barbells: *a sphere of radius, 10, with a cap removed and cylinders of radius, 5, of various lengths, $L = 1$, $L = 10$, $L = 30$, and $L = 50$, glued in, and then capped off with a hemisphere of radius 5.*

At the top, $L = 1$, so the data set lies on a shape which is almost a sphere and the first three nonconstant eigenfunctions are close to the spherical eigenfunctions, (X, Y, Z) that we discussed in Section 3.8. The eigenmap seems to be close to $(-Z(x_i), X(x_i), -Y(x_i))$ giving a rotated image of the original shape just as in Section 6.7. In the second row, $L = 10$, we still see an embedding but one

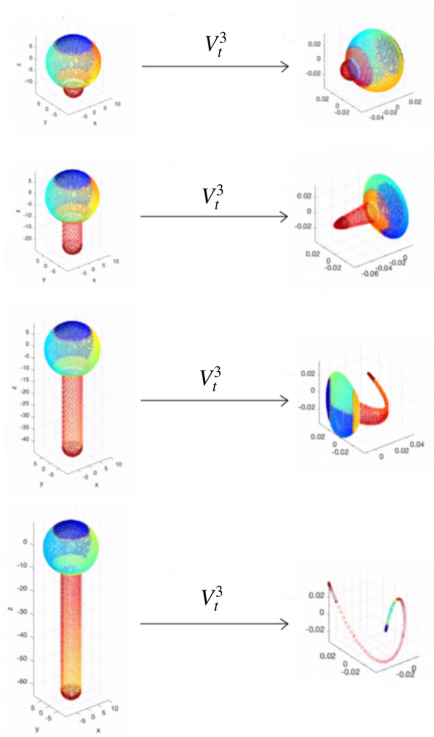


Figure 6.7: Here we see a data set of evenly distributed points on various lollipops mapped by an eigenmap, V_t^3 , which gives some possible insight into the image under a heat kernel embedding, ψ_t^3 of these lollipops. MATLAB computation by Maziar Farahzad, Julinda Pillati Mujo, and Esteban Alcantara.

direction is becoming more deformed. The rotational symmetry is preserved but changing the X axis to a Z axis.

In the third row, $L = 30$, and as with the long barbell, there is a first nonzero eigenvalue corresponding to the full height of the manifold which has an eigenfunction that looks something like $\cos(Z)$. The next two eigenfunctions seem to correspond to z and x . That is the points x_i are mapped to $(\cos(Z(x_i)), Z(x_i), X(x_i))$ and so the image is not an embedding. It is flattened and then bent around a cosine curve.

In the bottom row, $L = 100$ and now the fact that there is a sphere on one end seems to be undetectible. The first three eigenfunctions seem to all be functions

of z alone causing the image under the eigenmap to be a parametrized curve. This is similar to the effect that we saw with the thin torus in Section 3.7.

Please note again that these are images under the Belkin-Niyogi eigenmaps, V_t^3 not the images under the smooth eigenmaps, Φ_t^3 , nor the heat embeddings, Ψ_t^3 of Bérard-Besson-Gallot. It should be noted again that, to save on computation time, the students did not construct an adjacency graph but used the Euclidean distances to define their matrix W (since points that were not close lead to very small entries in W anyway).

6.10 Coifman and Lafon's Diffusion Maps of Data Sets

In 2004-2006 Coifman and Lafon extended this idea by introducing the notion of *diffusion maps* (DM) of high dimensional data sets [64][39] [38]. They allowed for data sets which are not uniformly distributed and discuss various kernel operators, P , based on the Riemannian Laplacian, based on the weighted graph Laplacian, and based on other linear operators. They studied the eigenvalues, $\tilde{\lambda}_j$, of these kernel's operators.

In particular, Coifman and Lafon studied the heat kernel of the Riemannian Laplacian, which they denoted as, $P_\tau = e^{-\tau\Delta}$. To understand this notation, we may use Taylor series,

$$P_\tau = e^{-\tau\Delta} = \sum_{n=0}^{\infty} \frac{(-\tau\Delta)^n}{n!} = I - \tau\Delta + \frac{(\tau\Delta)^2}{2!} - \frac{(\tau\Delta)^3}{3!} + \dots \quad (6.37)$$

Note that $e^{-\tau\Delta}$ has the same eigenfunctions, ϕ_j , as the Laplacian, $\Delta\phi_j = -\lambda_j\phi_j$, as seen by:

$$e^{-\tau\Delta}\phi_j = \sum_{n=0}^{\infty} \frac{(-\tau\Delta)^n}{n!}\phi_j = \sum_{n=0}^{\infty} \frac{(-\tau\lambda_j)^n}{n!}\phi_j = e^{-\lambda_j\tau}\phi_j. \quad (6.38)$$

Coifman and Lafon wrote their eigenvalues of P_τ as $\tilde{\lambda}_j^\tau$ decreasing from $\tilde{\lambda}_0^\tau = 1$ so that

$$P_\tau\phi_j = \tilde{\lambda}_j^\tau\phi_j. \quad (6.39)$$

When $P_\tau = e^{-\tau\Delta}$ we have

$$P_\tau\phi_j = \tilde{\lambda}_j^\tau\phi_j \iff \Delta\phi_j = -\lambda_j\phi_j \text{ so } \tilde{\lambda}_j = e^{-\lambda_j}. \quad (6.40)$$

Coifman-Lafon's *diffusion map* is a weighted eigenmap into Euclidean space:

$$\tilde{\Psi}_\tau^N(x) = \{\tilde{\lambda}_j^\tau\phi_j(x)\}_{j=1}^N \subset \mathbb{E}^N \quad (6.41)$$

Recall the *heat kernel embedding* of Bérard-Besson-Gallot we saw in (5.4) and observe that by taking $\tau = t/2$ and applying (6.40), Coifman-Lafon’s diffusion map is a truncation of the heat kernel embedding of Bérard-Besson-Gallot:

$$\tilde{\Psi}_{t/2}^N(x) = \{e^{-\lambda_j t/2} \phi_j(x)\}_{j=1}^N = \Psi_t^N(x). \tag{6.42}$$

They did not directly cite [16] but they do cite other papers and textbooks that cite the original article.

Coifman-Lafon demonstrated with graphics as in Figure 6.8 that this weighted eigenmap captures the intrinsic geometry of a data set lying on a curve. In particular they demonstrated that collections of points lying on a long twisted and knotted curve in the original space maps onto a perfectly round circle in \mathbb{E}^2 capturing the intrinsic geometry of the original curve. In [38], Coifman and Lafon proved rigorously how N must be large after τ is chosen to be small to justify the good behavior of their method for any member of their class of kernels.

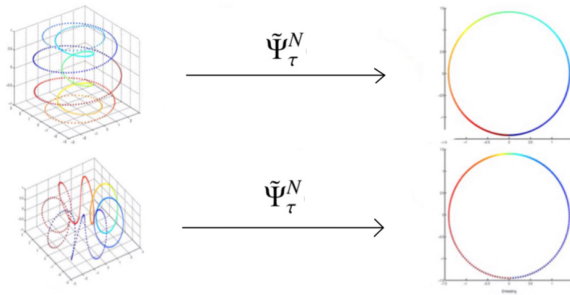


Figure 6.8: Coifman-Lafon demonstrate that the images of data sets of evenly distributed points lying on closed curves under a truncated diffusion map, $\tilde{\Psi}_\tau^N$, with $N=2$ is round circle. This embedded circle recovers the intrinsic Riemannian geometry of the 1 dimensional submanifolds containing the point sets. Image Credit: Excerpt from Figure 4 in [38].

6.11 Diffusion Distance

In [38], Coifman and Lafon coined the term **Diffusion Distance**

$$D_\tau(x, y)^2 = \sum_{j=1}^{\infty} \tilde{\lambda}_1^{2\tau} (\phi_j(x) - \phi_j(y))^2 = \sum_{j=1}^{\infty} e^{-\lambda_j t} (\phi_j(x) - \phi_j(y))^2 \tag{6.43}$$

which agrees with pull back of the ℓ^2 distance of Bérard-Besson-Gallot when applied to the heat kernel:

$$D_{t/2}(x, y) = \|\Psi_t(x) - \Psi_t(y)\|_{\ell^2} = (\Psi_t^* d_{\ell^2})(x, y). \quad (6.44)$$

While we have provided the relationship here between their work and that of Bérard-Besson-Gallot, remember that Coifman and Lafon's work applies to a larger class of kernels. Once again we apologize for not including probability references in this survey that inspired and helped justify the work of Coifman and Lafon.

It is important to note that the Diffusion Distance is an extrinsic distance measured using direct line segments in ℓ^2 . Even when studying the heat kernel,

$$D_{t/2}(x, y)^2 = \sum_{j=1}^{\infty} e^{-\lambda_j t} (\phi_j(x) - \phi_j(y))^2. \quad (6.45)$$

is an extrinsic distance between points and does not approximate $d_g(x, y)$. Even with appropriate scaling and taking $t \rightarrow 0$, the heat kernel map is not distance preserving, but is only length preserving as we discussed in Section 5.7. Nevertheless $D_{t/2}(x, y)$ is an interesting distance function on the original manifold M . It induces rectifiable lengths of curves, $C : [0, 1] \rightarrow M$,

$$L_t(C) = \sup \left\{ \sum_{i=1}^n D_{t/2}(C(s_{i-1}), C(s_i)) : 0 \leq s_0 < s_1 < \dots < s_n = 1 \right\}, \quad (6.46)$$

where the sup is taken over all partitions, $0 \leq s_0 < s_1 < \dots < s_n = 1$. These lengths agree with the lengths of curves defined using the pullback of the Euclidean metric tensor. By the work of Bérard-Besson-Gallot in [16], with appropriate rescaling, these lengths converge

$$\lim_{t \rightarrow 0} L_t(C) = L_g(C). \quad (6.47)$$

With that rescaling we have (5.52).

Memoli and Sapiro computed the intrinsic distance between points in data sets lying on a manifold using related ideas in [75].

7 Spectral Embeddings and Embedding Dimensions

In this section we present various results on *spectral embeddings* of Riemannian manifolds where we refer to an embedding as spectral if it is defined using

eigenfunctions, eigenvalues, and/or heat kernels. There has been significant research activity in this area inspired by the applications of spectral embeddings to study high dimensional data sets. It is of particular interest to study the smallest dimension, N , such that M^n can be embedded into \mathbb{R}^N and to find uniform bounds on this dimension N . This is referred to as the *embedding dimension* of M^n .

Note a truncated heat kernel map, ψ_t^N , as defined in (5.12) is an embedding iff the corresponding smooth eigenmap, Φ^N , as defined in (6.7) is also an embedding. This is because the only distinction between these maps is the scaling of the various entries using eigenvalues. Other spectral maps discussed above have similar uniform bounds on N .

7.1 Deforming Manifolds and their Embeddings

As we saw in Section 5, Bérard, Besson, and Gallot proved that for any $\epsilon > 0$ there is a $t_\epsilon > 0$ sufficiently small such that for any $t \in (0, t_\epsilon)$, there is an $N_{t,\epsilon}$ sufficiently large that for any $N > N_{t,\epsilon}$ we have

$$|(\psi_t^N)^* g_{\mathbb{E}^N} - g| < \epsilon \quad (7.1)$$

where

$$\psi_t^N(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \{e^{-\lambda_j t/2} \phi_j(x)\}_{j=1}^N \in \mathbb{E}^N. \quad (7.2)$$

In particular for t sufficiently small, and N sufficiently large, they can guarantee that $(\psi_t^N)^* g_{\mathbb{E}^N}$ is close enough to g to be positive definite. In fact for sufficiently large N the truncated eigenmap, ψ_t^N , is an embedding. We saw that for a thin torus, we would need to take N very large to achieve this in Section 5.7. We also saw that the extrinsic distances were not well controlled as $t \rightarrow 0$ in Section 5.7.

If the manifold has computable eigenfunctions, like a sphere or a torus, then we can compute t_ϵ and $N_{t,\epsilon}$ for that specific manifold. In fact, we know we only need three eigenfunctions on a sphere and four on a torus to achieve an embedding. Naturally it is of interest to have similarly strong controls on manifolds that are close to these special manifolds.

In fact Bérard, Besson, and Gallot proved their embeddings are stable under C^0 perturbations of the metric with a uniform lower bound on Ricci curvature (See Theorem 21 in [16]):

Theorem 7.1. [16] Let (M, g) be a closed n -dimensional Riemannian manifold, $\epsilon_0 > 0$ and N_0 be a positive integer. Let g' be any metric on M such that

$$(1 - \epsilon)g \leq g' \leq (1 + \epsilon)g, \epsilon < \epsilon_0. \quad (7.3)$$

We assume that all metrics under consideration satisfy $\text{Ric}_{(M,g')} \geq -(n-1)H$ for some constant $H \geq 0$. Then there exist constants $\eta_{g,i,H}(\epsilon)$, $1 \leq i \leq N_0$, which go to 0 with ϵ , such that to any orthonormal basis $\{\phi'_i\}$ of eigenfunctions of $\Delta_{g'}$ we can associate an orthonormal basis $\{\phi_i\}$ of eigenfunctions of Δ_g satisfying

$$\|\phi_i - \phi'_i\|_{L^\infty} \leq \eta_{g,i,H}(\epsilon) \quad (7.4)$$

for all $i \leq N_0$.

As a consequence, suppose we consider the standard sphere \mathbb{S}^2 . We saw in Section 5.4 that its first three nonconstant eigenfunctions map it to a rescaled sphere in \mathbb{E}^3 . If (M, g) is C^2 close to \mathbb{S}^2 then its eigenfunctions are close to that of the standard sphere and its truncated eigenmap will map it close to a sphere. If (M, g) is only C^0 close to \mathbb{S}^2 but has Ricci curvature bounded below, then this works as well. This works for lollipops with fixed cylinder radius, r , and length $L_i \rightarrow 0$. The increasingly thin lollipops with $r_i \rightarrow 0$ and L fixed are not C^0 converging to a standard sphere and do not have Ricci curvature bounded from below uniformly, so we cannot apply the Bérard, Besson, and Gallot C^0 stability theorem to study them.

7.2 Local Embeddings of M^n into \mathbb{E}^n

In 2008, Jones, Maggioni, and Schul studied local properties of embeddings using eigenfunctions. They showed that any smooth n -dimensional manifold admitting charts in which the metric is C^α can be locally embedded in \mathbb{E}^n by eigenfunctions of the Laplace operator [58, 59]. These coordinates are bi-Lipschitz on embedded balls of the domain or manifold, with distortion constants that depend only on natural geometric properties of the domain or manifold.

7.3 Embedding Dimensions of Spectral Embeddings

In 2014 Bates considered the class

$$\mathcal{M}_{H,i_0}^V = \{(M, g) : \dim M = n, \text{Ric} \geq -(n-1)H, \text{inj}(M) \geq i_0, \text{Vol}(M) \leq V\} \quad (7.5)$$

in [10]. Recall this is the class Anderson-Cheeger proved was C^α compact in [8].

Bates proved that any closed, connected Riemannian manifold M can be embedded using a smooth truncated eigenmap where the number of required eigenfunctions is uniform on the class:

Theorem 7.2. [10] Given any $H \in \mathbb{R}$, any $i_0 > 0$ and any $V > 0$, there is a positive integer $N = N_{H,n,i_0,V}$ and constant $\epsilon_{H,n,i_0,V} > 0$, such that, for any $M^n \in \mathcal{M}_{H,i_0}^V$ and for all $z \in M$,

$$\Phi^N : M \rightarrow \mathbb{E}^N \text{ where } x \mapsto (\phi_1(x), \dots, \phi_N(x)) \tag{7.6}$$

is a smooth embedding.

As a consequence there is a uniform upper bound on the embedding dimension of M on this class using eigenmaps as embeddings. Note that in his proof, Bates applied the work of Jones, Maggioni, and Schul that we mentioned above.

7.4 Uniform Bounds on Spectral Embeddings by Portegies

In 2016, Portegies considered various spectral embeddings on the same class \mathcal{M}_{H,i_0}^V . In particular he studied the truncated heat kernel map of Bérard-Besson-Gallot,

$$\psi_t^N(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \{e^{-\lambda_j t/2} \phi_j(x)\}_{j=1}^N \in \mathbb{E}^N, \tag{7.7}$$

that we reviewed in Section 5.2. Recall that Bérard-Besson-Gallot had proven that the pull back of the metric tensor under their map is close to the original metric tensor under the following limits:

$$\lim_{t \rightarrow 0^+} \lim_{N \rightarrow \infty} (\psi_t^N)^* g_{\mathbb{E}^N}(V_x, V_x) = g(V_x, V_x). \tag{7.8}$$

Portegies proved this convergence is uniform on \mathcal{M}_{H,i_0}^V in [81]:

Theorem 7.3. Let $\epsilon > 0$. Then there exists a $t_0 = t_0(n, \epsilon, H, i_0)$ such that for all $0 < t \leq t_0$ there exists an $N_0 = N_0(n, \epsilon, t, H, i_0, V)$ such that if $N \geq N_0$, for all $(M, g) \in \mathcal{M}_{H,i_0}^V$, the map

$$\psi_t^N(x) = (2t)^{(n+2)/4} \sqrt{2}(4\pi)^{n/4} (e^{-\lambda_1 t} \phi_1(x), \dots, e^{-\lambda_N t} \phi_N(x)) \tag{7.9}$$

is an embedding of M into \mathbb{E}^N such that

$$1 - \epsilon < |(d\psi_t^N)_p| < 1 + \epsilon. \tag{7.10}$$

In particular, the sequence of Riemannian manifolds $(M, (\psi_t^N)^* g_{\mathbb{E}^N})$ converges in the C^0 sense as $N \rightarrow \infty$ to the original manifold (M, g) . The speed of this convergence is the same for any $(M, g) \in \mathcal{M}_{H,i_0}^V$. Note that the sequence of Riemannian manifolds $(M, (\Phi^N)^* g_{\mathbb{E}^N})$ defined by pulling back metric tensors

using eigenmaps diverges. We need the rescaling by the eigenvalues prevent the lengths from stretching to infinity in the limit.

Portegies proved this theorem by first obtaining a quantitative version of the estimate on the harmonic radius found by Anderson-Cheeger in [8]. He then studied the images under the map ψ_t^N restricted to these harmonic charts. In addition to proving the above theorem he proved bounds on other truncated spectral embeddings defined using heat kernels [81].

7.5 Spectral Embeddings depending on the Connection Laplacian

There are other spectral embeddings based on other types of Laplacian operators, such as the connection Laplacian. The connection Laplacian is a Laplace operator acting on the various tensor bundles of a manifold. Inspired by the cryo-electron microscopy image problem for reconstructing macromolecular structures, Singer-Wu [85] introduced the *vector diffusion maps* (VDM). The VDM method helped them with aligning images and finding similar images (nearest neighbors) more effectively than the diffusion maps studied by Coifman-Lafon.

The VDM is a spectral embedding [70] that depends on the connection Laplacian associated with a possibly nontrivial bundle structure. Its general goal is to integrate local information and the relationship between these pieces of local information in order to obtain the global information of the dataset; for example, the ptychographic imaging problem [73], the synchronization problem [9], the vector nonlocal mean/median [69], the orientability problem [87], etc. Numerically, the VDM depends on the spectral study of the graph connection Laplacian [33, 50, 51, 85, 86], which is a direct generalization of the graph Laplacian discussed in the spectral graph theory text by Chung [31].

In [70], Lin and Wu investigated estimates of eigenvector fields and truncations of heat kernels and its derivatives. In the end, they showed that an embedding of a manifold can be constructed with finite eigenvector fields of the connection Laplacian for a class of closed smooth manifolds

$$\mathcal{M}_{n,H,i_0,V} = \{(M, g) : |\text{Ric}_g| \leq H, \text{inj}(M) \geq i_0, \text{Vol}(M) \leq V\}. \quad (7.11)$$

Recall this is the class that Anderson proved was $C^{1,\alpha}$ precompact in [7].

In [71], Lin proved that, in addition, we can ensure the pull back of the metric tensor is close to the original metric tensor under a limiting process and this can be done uniformly on the class $\mathcal{M}_{n,H,i_0,V}^H$ obtaining the following theorem:

Theorem 7.4. [71] Given $\epsilon > 0$, there exists $t_0 = t_0(n, H, i_0, \epsilon)$ so that for all $t < t_0$ there exists $N_0 = N_0(n, H, i_0, \epsilon, t, V)$ such that for all $(M, g) \in \mathcal{M}_{n,H,i_0,V}$,

there exist points p_1, \dots, p_{N_0} on M such that the map $\mathcal{H} : (M^n, g) \rightarrow \mathbb{E}^{N_0}$ defined by

$$\mathcal{H}(p) = \frac{A(2t)^{(3n+2)/4}}{V_e} \left(\|K_{TM}(p, t; p_1)\|_{HS}^2, \dots, \|K_{TM}(p, t; p_{N_0})\|_{HS}^2 \right) \quad (7.12)$$

is an embedding, where $A = A(n, H, i_0, \epsilon, t, V)$ is a constant, K_{TM} denotes the heat kernel of the connection Laplacian on the tangent bundle TM , $\|\cdot\|_{HS}$ denotes the Hilbert-Schmidt norm, and

$$V_e = \left(\int_{\mathbb{R}^n} (\partial_{x_1} \|K_{TM}^{\mathbb{E}^n}(0, \frac{1}{2}; y)\|_{HS}^2)^2 dy \right)^{\frac{1}{2}}.$$

Here, $K_{TM}^{\mathbb{E}^n}$ is the standard Euclidean heat kernel on \mathbb{E}^n . In addition,

$$1 - \epsilon < |(d\mathcal{H})_p| < 1 + \epsilon \text{ for all } p \in M.$$

The same statements hold with every heat kernel K_{TM} replaced by the truncated version K_{TM}^N . See [71].

The proof of this theorem also relies on a quantitative version of the estimate on the harmonic. Since Lin must control the tangent bundle TM and the heat kernel of the connection Laplacian action on sections of the tangent bundle in [71], it is not surprising that Lin needs to study a class with $C^{1,\alpha}$ compactness.

7.6 L^p Bounds on Pull Back Metrics

In 2021, Ambrosio, Honda, Portegies and Tewodrose considered the class

$$\mathcal{M}_{H,V}^D = \{(M, g) : |\text{Ric}_g| \leq H, \text{Vol}(M) \geq V, \text{diam}(M) \leq D\}. \quad (7.13)$$

in [6]. Note that this class has no uniform lower bound on injectivity radius. This is the class of noncollapsing manifolds with lower Ricci curvature bounds studied by Cheeger-Colding that have subsequences which converge in the metric measure sense to limits and have controls on their eigenfunctions [25]. Yu Ding controlled the heat kernels on such spaces in [45]. The limits of these spaces are in fact $\text{RCD}^*(H, N)$ spaces which were first introduced by Ambrosio-Gigli-Savare [5].

In fact Ambrosio, Honda, Portegies and Tewodrose considered spectral maps defined on these $\text{RCD}^*(H, N)$ spaces in [6] extending many known results to this class. Theorem 6.8 and 6.9 in their paper [6] are also new results for smooth Riemannian manifolds in $\mathcal{M}_{H,V}^D$. Theorem 6.9 concerns an eigenmap, so we include it below as reformulated for us by Portegies for manifolds:

Theorem 7.5 (Theorem 6.9 in [6]). For all $D > 0, V > 0, \epsilon > 0$ and $1 \leq p < \infty$ there exists

$$t_0 := t_0(H, N, V, D, \epsilon, p) > 0$$

such that for all $0 < t \leq t_0$ and every closed n -dimensional Riemannian manifold $M \in \mathcal{M}_{H,V}^D$, there exists an $N_0 = N_0(H, N, V, D, \epsilon, p, t)$ such that for all $N \geq N_0$, we have

$$\| |\omega_{nt}^{(n+2)/2} g_t^N - c_n g|_{HS} \|_{L^p} \leq \epsilon.$$

where

$$g_t^N := \Psi_t^* g_{\mathbb{E}^N} = \sum_{i=1}^N e^{-2\lambda_i t} d\phi_i \otimes d\phi_i.$$

Observe that this theorem is stating that the pull back metric tensor is converging to the original tensor but only with uniform L^p bounds on the convergence. In Conjectures 8.4 and 8.5 stated at the end of this paper, we ask whether we can achieve Gromov-Hausdorff or Intrinsic Flat Convergence with the right choice of class \mathcal{M} of Riemannian manifolds. Note that L^p convergence of the metric tensor is not enough to achieve uniform, Gromov-Hausdorff, or intrinsic flat convergence unless there are also uniform bounds on the metric tensors from above and below as shown by Allen-Sormani in [3] and [4] or controls like those required by Allen-Perales-Sormani to achieve VADB convergence in [2].

8 Open Questions

Throughout this paper we have been exploring the properties of heat kernels, eigenfunctions, and spectral maps for a variety of manifolds and classes of manifolds. In this section, we vaguely state some conjectures related to these ideas. We hope to bring together mathematicians from around the world to study these problems together.

8.1 Conjectures on Truncated Eigenmaps of Small Dimension

In the work of Belkin-Niyogi reviewed in Section 6.1 and the work of Coifman-Lafon reviewed in Section 6.10, truncated spectral maps were quite useful even when $N = 2$ or $N = 3$ and without taking $t \rightarrow 0$. This is due in part to the underlying symmetries of the data sets that they are studying. In Section 5.4 we saw that the symmetries of a sphere guarantee that it spectrally embeds into Euclidean space with only $N = 3$ and does so symmetrically. Data sets lying on surfaces which are close to a sphere in the appropriate sense also

have nearby eigenfunctions and thus also embed with $N = 3$. Other surfaces like the barbell and the lollipop studied by our students did not always spectrally embed with $N = 3$ as seen in Sections 6.8 and 6.9. Even the torus $\mathbb{S}^1 \times \mathbb{S}^1$ needs $N = 4$ and a thin torus needs a much larger N as discussed in Sections 5.5 and 5.6.

This leads naturally to the following conjecture:

Conjecture 8.1. Can we describe a class of manifolds \mathcal{M} for which

$$\Psi_t^N : M \rightarrow \mathbb{E}^N \tag{8.1}$$

is an embedding when $N = 3$ or $N = 4$?

Since eigenvalues and eigenfunctions vary smoothly with C^2 convergence of manifolds, we know that this conjecture holds with $N = 3$ for the class of manifolds which are sufficiently close to a standard sphere in the C^2 sense, and it holds with $N = 4$ for the class of manifolds which are sufficiently close to a standard torus in the C^2 sense. We saw in Section 7.1 that Bèard-Besson-Gallot proved that if M have Ricci curvature bounded uniformly from below and are C^0 close, then their eigenfunctions and eigenvalues are close, thus this conjecture holds with $N = 3$ for classes of manifolds C^0 close to the standard sphere with Ricci curvature bounded uniformly from below, and similarly for $N = 4$ near the standard torus. However there is no precise statement as to how close these manifolds must be. It would also be interesting to define a class which does not mention the sphere or the torus.

8.2 Conjectures on Selective Eigenmaps

In Section 5.6 we observed how shapes like a thin torus would not be embedded by a truncated spectral map until N is taken very large and then there has not been a very useful reduction in dimension. We suggested the idea of selecting eigenfunctions to define an embedding rather than just truncating a spectral embedding. In particular, we suggested only selecting eigenfunctions with eigenvalues near a given eigenvalue λ . This leads to the following natural conjecture:

Conjecture 8.2. If we are given a Riemannian manifold, M^n , is a class, \mathcal{M} , with $\text{injrad} \geq i_0 > 0$ and we define a *selected spectral map* using only eigenfunctions whose eigenvalues are close to λ_0 defined as a function of i_0 , can we prove this selected spectral map is an spectral embedding?

Note that in the conjecture above we have been deliberately vague about the choice of λ_0 and how it depends on i_0 . We have also been vague about the class, \mathcal{M} . It might be natural to require the class have bounds on Ricci curvature, volume, or diameter.

Conjecture 8.3. Can we describe a class of manifolds \mathcal{M} for which we can find *selected spectral map* that is an embedding when $N = 3$ or $N = 4$?

8.3 Conjectures on Convergence of Pullbacks of Embeddings

Recall in Section 5.2 we reviewed the truncated heat kernel map of Bérard-Besson-Gallot,

$$\psi_t^N(x) = \sqrt{2}(4\pi)^{n/4} t^{(n+2)/4} \{e^{-\lambda_j t/2} \phi_j(x)\}_{j=1}^N \in \mathbb{B}^N, \quad (8.2)$$

has the property that: $\forall \epsilon > 0, \exists t_\epsilon > 0, \text{ s.t. } \forall t \in (0, t_\epsilon), \exists N_{t,\epsilon} \text{ s.t. } \forall N \geq N_t$

$$|g_t^N - g| < \epsilon \quad (8.3)$$

where $g_t^N = (\psi_t^N)^* g_{\mathbb{B}^N}$. In Section 7.4, we saw that Portegies proved the choice of t_ϵ and of $N_{t,\epsilon}$ can be made uniform on the class \mathcal{M}_{H,i_0}^V .

Recalling the definition of the Lipschitz distance in Section 4.4, we can rewrite this as $\forall \epsilon > 0, \exists t_\epsilon > 0, \text{ s.t. } \forall t \in (0, t_\epsilon), \exists N_{t,\epsilon} \text{ s.t. } \forall N \geq N_t$

$$d_{Lip}((M, d_{g_t^N}) - (M, d_g)) < \epsilon \quad (8.4)$$

where the choice of t_ϵ and of $N_{t,\epsilon}$ can be made uniform on the class \mathcal{M}_{H,i_0}^V .

Recall the definition of Gromov-Hausdorff distance in Section 4.9.

Conjecture 8.4. Can we find a class of manifolds, \mathcal{M} such that $\forall \epsilon > 0, \exists t_\epsilon > 0, \text{ s.t. } \forall t \in (0, t_\epsilon), \exists N_{t,\epsilon} \text{ s.t. } \forall N \geq N_t$

$$d_{GH}((M, d_{g_t^N}) - (M, d_g)) < \epsilon \quad (8.5)$$

where the choice of t_ϵ and of $N_{t,\epsilon}$ can be made uniform on the class \mathcal{M} ?

Recall the definition of the intrinsic flat distance in Section 4.13.

Conjecture 8.5. Can we find a class of manifolds, \mathcal{M} such that $\forall \epsilon > 0, \exists t_\epsilon > 0, \text{ s.t. } \forall t \in (0, t_\epsilon), \exists N_{t,\epsilon} \text{ s.t. } \forall N \geq N_t$

$$d_{GH}((M, d_{g_t^N}) - (M, d_g)) < \epsilon \quad (8.6)$$

where the choice of t_ϵ and of $N_{t,\epsilon}$ can be made uniform on the class \mathcal{M} ?

Since Lipschitz convergence implies Gromov-Hausdorff and Intrinsic Flat convergence, both of the above conjectures are immediately true for the class \mathcal{M}_{H,i_0}^V . However they should be provable without the injectivity radius bound. See the discussion in Section 7.6.

8.4 Conjectures on Convergence of Manifolds and their Embeddings

We saw in Section 7.1 that Bèrard-Besson-Gallot proved that if M_j have Ricci curvature bounded uniformly from below and $M_j \rightarrow M_\infty$ in the C^0 sense, then $\psi_t^N(M_j) \rightarrow \psi_t^3(M_\infty)$ in the C^0 sense because the eigenfunctions converge.

Recall the definition of Hausdorff and Gromov-Hausdorff convergence in Sections 4.9 and 4.11:

Conjecture 8.6. Suppose $M_j \subset \mathcal{M}$ for some class \mathcal{M} , and $M_j \rightarrow M_\infty$ in the measured Gromov-Hausdorff sense, does $\psi_t^N(M_j) \rightarrow \psi_t^N(M_\infty)$ in the Hausdorff sense as subsets of Euclidean Space, \mathbb{E}^N ?

Recall the definition of the intrinsic flat convergence in Section 4.13.

Conjecture 8.7. Suppose $M_j \subset \mathcal{M}$ for some class \mathcal{M} , and $M_j \rightarrow M_\infty$ in the SWIF sense, does $\psi_t^N(M_j) \rightarrow \psi_t^N(M_\infty)$ in the Flat sense as submanifolds of Euclidean Space, \mathbb{E}^N ?

These conjectures are quite deliberately vague as the specific class of manifolds has not been prescribed. Both conjectures are probably easily proven to be true if we choose a class with a lower bound on injectivity radius and on Ricci curvature using the theorems surveyed within this article. It would be far more challenging to remove the injectivity radius bound.

Please contact us if you would like to work on any of the conjectures in this survey. We are happy to set up teams of collaborators to work together and might invite you to participate in workshops on the topic as well.

Acknowledgements

We would particularly like to thank our 2020 student research team members: Esteban Alcantara, Julinda Pillati Mujo, and Maziar Farahzad. Professor Sormani would like to thank the organizers of the 2022 TWAS TYAN Virtual Workshop on Differential Geometry for the honor of being invited to moderate a

session. It was truly inspiring to meet Professor Romain Murenzi, to hear him speak of his experience as a mathematics student and teacher from Burundi, and how important it can be to work in a young active subject. He spoke of how Professor Alex Grossman adopted him into his research group and gave him the opportunity to speak internationally, and encouraged us all to do the same for young scientists from developing countries. Here at the City University of New York, Professor Chen-Yun Lin and Professor Sormani often teach first and second generation students from developing countries and are delighted to build connections with faculty and students internationally.

References

- [1] Hiba Abdalla. Embedding Riemannian manifolds via their eigenfunctions and their heat kernel. *Bulletin of the Korean Mathematical Society*, 49(5):939–947, 2012.
- [2] Brian Allen, Raquel Perales, and Christina Sormani. Volume above distance below. *Journal of Differential Geometry*, 2022.
- [3] Brian Allen and Christina Sormani. Contrasting various notions of convergence in geometric analysis. *Pacific Journal of Mathematics*, 303(1):1–46, 2019.
- [4] Brian Allen and Christina Sormani. Relating notions of convergence in geometric analysis. *Nonlinear Analysis*, 200, 2020.
- [5] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré. Metric measure spaces with riemannian ricci curvature bounded from below. *Duke Math. J.*, 163(7):1405–1490, 2014.
- [6] Luigi Ambrosio, Shouhei Honda, Jacobus W Portegies, and David Tewodrose. Embedding of $rcd^*(k, n)$ spaces in l^2 via eigenfunctions. *Journal of Functional Analysis*, 280(10):108968, 2021.
- [7] Michael T. Anderson. Convergence and rigidity of manifolds under ricci curvature bounds. *Inventiones mathematicae*, 102(1):429–445, 1990.
- [8] Michael T. Anderson and Jeff Cheeger. C^α -compactness for manifolds with Ricci curvature and injectivity radius bounded below. *Journal of Differential Geometry*, 35(2):265 – 281, 1992.

- [9] Afonso S. Bandeira, Amit Singer, and Daniel A. Spielman. A Cheeger inequality for the graph connection Laplacian. *SIAM J. Matrix Anal. Appl.*, 34(4):1611–1630, 2013.
- [10] Jonathan Bates. The embedding dimension of laplacian eigenfunction maps. *Applied and Computational Harmonic Analysis*, 37(3):516 – 530, 2014.
- [11] M. Belkin and P. Niyogi. Towards a theoretical foundation for Laplacian-based manifold methods. In *Proceedings of the 18th Conference on Learning Theory (COLT)*, pages 486–500, 2005.
- [12] Mikhail Belkin. Problems of learning on manifolds. *Doctoral Thesis*, 2003.
- [13] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps and spectral techniques for embedding and clustering. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, NIPS’01, pages 585–591, Cambridge, MA, USA, 2001. MIT Press.
- [14] Mikhail Belkin and Partha Niyogi. Convergence of laplacian eigenmaps. *Advances in neural information processing systems*, 19, 2006.
- [15] Mikhail Belkin, Partha Niyogi, and Vikas Sindhwani. Manifold regularization: A geometric framework for learning from labeled and unlabeled examples. *Journal of machine learning research*, 7(11), 2006.
- [16] Pierre Bérard, G. Besson, and Sylvestre Gallot. Embedding riemannian manifolds by their heat kernel. *Geometric and Functional Analysis*, 4:373–398, 07 1994.
- [17] Marcel Berger, Paul Gauduchon, and Edmond Mazet. *Le spectre d’une variété riemannienne*. Lecture Notes in Mathematics, Vol. 194. Springer-Verlag, Berlin-New York, 1971.
- [18] Dmitri Burago, Yuri Burago, and Sergei Ivanov. *A course in metric geometry*, volume 33 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2001.
- [19] Dmitri Burago, Sergei Ivanov, and Yaroslav Kurylev. A graph discretization of the laplace–beltrami operator. *Journal of Spectral Theory*, 4(4):675–714, 2015.

- [20] Peter Buser. On cheeger's inequality $\lambda_1 \geq h/4$. In *AMS Proc. Symposia Pure Maths*, volume 36, pages 29–78, 1980.
- [21] Peter Buser. A note on the isoperimetric constant. *Ann. Sci. École Norm. Sup. (4)*, 15(2):213–230, 1982.
- [22] Isaac Chavel. *Eigenvalues in Riemannian geometry*. Academic Press Orlando, 1984.
- [23] Jeff Cheeger. *Comparison and finiteness theorems for Riemannian manifolds*. Princeton University, 1967.
- [24] Jeff Cheeger. Finiteness theorems for Riemannian manifolds. *Amer. J. Math.*, 92:61–74, 1970.
- [25] Jeff Cheeger and Tobias H. Colding. On the structure of spaces with Ricci curvature bounded below. III. *J. Differential Geom.*, 54(1):37–74, 2000.
- [26] Jeff Cheeger, Mikhail Gromov, and Michael Taylor. Finite propagation speed, kernel estimates for functions of the laplace operator, and the geometry of complete riemannian manifolds. *Journal of Differential Geometry*, 17(1):15–53, 1982.
- [27] Jeff Cheeger and Shing-Tung Yau. A lower bound for the heat kernel. *Communications on Pure and Applied Mathematics*, 34(4):465–480, 1981.
- [28] Shiu-Yuen Cheng. Eigenvalue comparison theorems and its geometric applications. *Mathematische Zeitschrift*, 143(3):289–297, 1975.
- [29] Shiu-Yuen Cheng and Peter Li. Heat kernel estimates and lower bound of eigenvalues. *Commentarii Mathematici Helvetici*, 56(1):327–338, 1981.
- [30] Siu Yuen Cheng, Peter Li, and Shing-Tung Yau. On the upper estimate of the heat kernel of a complete riemannian manifold. *American Journal of Mathematics*, 103(5):1021–1063, 1981.
- [31] F. Chung. *Spectral Graph Theory*. American Mathematical Society, 1996.
- [32] F. Chung, Alexander Grigorian, and S.-T Yau. Upper bounds for eigenvalues of the discrete and continuous laplace operators. *Advances in Mathematics*, 117:165–178, 02 1996.

- [33] F Chung, W Zhao, and M Kempton. Ranking and sparsifying a connection graph. In Anthony Bonato and Jeannette Janssen, editors, *Algorithms and Models for the Web Graph*, volume 7323 of *Lecture Notes in Computer Science*, pages 66–77. Springer Berlin Heidelberg, 2012.
- [34] F. R. K. Chung. Diameters and eigenvalues. *J. Amer. Math. Soc.*, 2(2):187–196, 1989.
- [35] F. R. K. Chung and S.-T. Yau. Eigenvalues of graphs and sobolev inequalities. *Combinatorics, Probability and Computing*, 4(1):11–25, 1995.
- [36] F. R. K. Chung and S.-T. Yau. A combinatorial trace formula. In *Tsing Hua lectures on geometry & analysis (Hsinchu, 1990–1991)*, pages 107–116. Int. Press, Cambridge, MA, 1997.
- [37] Fan Chung. The heat kernel as the pagerank of a graph. *Proceedings of the National Academy of Sciences*, 104(50):19735–19740, 2007.
- [38] Ronald R. Coifman and Stéphane Lafon. Diffusion maps. *Applied and Computational Harmonic Analysis*, 21(1):5 – 30, 2006. Special Issue: Diffusion Maps and Wavelets.
- [39] Ronald R Coifman, Stephane Lafon, Ann B Lee, Mauro Maggioni, Boaz Nadler, Frederick Warner, and Steven W Zucker. Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps. *Proceedings of the national academy of sciences*, 102(21):7426–7431, 2005.
- [40] R. Courant, K. Friedrichs, and H. Lewy. Über die partiellen Differenzengleichungen der mathematischen Physik. *Math. Ann.*, 100(1):32–74, 1928.
- [41] Mandouvalos N. Davies E.B. Heat kernel bounds on hyperbolic space and kleinian groups. *Proc. London Math. Soc.*(3), 52(1):182–208, 1988.
- [42] Amédée Debiard, Bernard Gaveau, and Edmond Mazet. Théoremes de comparaison en géométrie riemannienne. *Publications of the Research Institute for Mathematical Sciences*, 12(2):391–425, 1976.
- [43] Dennis DeTurck and Jerry L. Kazdan. Some regularity theorems in riemannian geometry. *Annales scientifiques de l'École Normale Supérieure*, Ser. 4, 14(3):249–260, 1981.

- [44] E. W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.
- [45] Yu Ding. Heat kernels and Green's functions on limit spaces. *Comm. Anal. Geom.*, 10(3):475–514, 2002.
- [46] Manfredo Perdigão do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.
- [47] Jozef Dodziuk. Difference equations, isoperimetric inequality and transience of certain random walks. *Trans. Amer. Math. Soc.*, 284:787–794, 02 1984.
- [48] Jozef Dodziuk and Leon Karp. Spectral and function theory for combinatorial Laplacians. In *Geometry of random motion (Ithaca, N.Y., 1987)*, volume 73 of *Contemp. Math.*, pages 25–40. Amer. Math. Soc., Providence, RI, 1988.
- [49] Edwards. *Riemann's Zeta Function*. Pure and Applied Mathematics. Academic Press, 1974.
- [50] N El Karoui and H.-T. Wu. Connection graph Laplacian and random matrices with random blocks. *Information and Inference: A Journal of the IMA*, 4:1–42, 2015.
- [51] Nouredine El Karoui, Hau-Tieng Wu, et al. Graph connection laplacian methods can be made robust to noise. *The Annals of Statistics*, 44(1):346–372, 2016.
- [52] Koji Fujiwara. Eigenvalues of laplacians on a closed riemannian manifold and its nets. *Proceedings of the American Mathematical Society*, 123(8):2585–2594, 1995.
- [53] Kenji Fukaya. Collapsing of riemannian manifolds and eigenvalues of laplace operator. *Inventiones mathematicae*, 87(3):517–547, 1987.
- [54] Alexander Grigor'yan and Masakazu Noguchi. The heat kernel on hyperbolic space. *Bull. London Math. Soc.*, 30:643–650, 1998.
- [55] Mikhael Gromov. *Structures métriques pour les variétés riemanniennes*, volume 1 of *Textes Mathématiques [Mathematical Texts]*. CEDIC, Paris, 1981. Edited by J. Lafontaine and P. Pansu.

- [56] M. Hein, J. Audibert, and U. von Luxburg. From graphs to manifolds - weak and strong pointwise consistency of graph Laplacians. In *Proceedings of the 18th Conference on Learning Theory (COLT)*, pages 470–485, 2005.
- [57] H Hess, R Schrader, and DA Uhlenbrock. Kato’s inequality and the spectral distribution of laplacians on compact riemannian manifolds. *Journal of Differential Geometry*, 15(1):27–37, 1980.
- [58] Peter W. Jones, Mauro Maggioni, and Raanan Schul. Manifold parametrizations by eigenfunctions of the Laplacian and heat kernels. *Proc. Natl. Acad. Sci. USA*, 105(6):1803–1808, 2008.
- [59] Peter W Jones, Mauro Maggioni, and Raanan Schul. Universal local parametrizations via heat kernels and eigenfunctions of the laplacian. *Ann. Acad. Sci. Fenn. Math.*, 35:131–174, 2010.
- [60] Atsushi Kasue. Convergence of Riemannian manifolds and Laplace operators. I. *Ann. Inst. Fourier (Grenoble)*, 52(4):1219–1257, 2002.
- [61] Atsushi Kasue and Hironori Kumura. Spectral convergence of riemannian manifolds. *Tohoku Mathematical Journal, Second Series*, 46(2):147–179, 1994.
- [62] Atsushi Kasue and Hironori Kumura. Spectral convergence of conformally immersed surfaces with bounded mean curvature. *The Journal of Geometric Analysis*, 12(4):663–681, 2002.
- [63] Risi Imre Kondor and John Lafferty. Diffusion kernels on graphs and other discrete structures. In *Proceedings of the 19th international conference on machine learning*, volume 2002, pages 315–322, 2002.
- [64] Stéphane S Lafon. *Diffusion maps and geometric harmonics*. Yale University, 2004.
- [65] Gregory F. Lawler and Alan D. Sokal. Bounds on the L^2 spectrum for Markov chains and Markov processes: a generalization of Cheeger’s inequality. *Trans. Amer. Math. Soc.*, 309(2):557–580, 1988.
- [66] Peter Li. Large time behavior of the heat equation on complete manifolds with non-negative ricci curvature. *Annals of mathematics*, 124(1):1–21, 1986.

- [67] Peter Li and Shing Tung Yau. On the parabolic kernel of the schrödinger operator. *Acta Mathematica*, 156:153–201, 1986.
- [68] Peter Li and S.T. Yau. Estimates of eigenvalues of a compact riemannian manifold. In *Geometry of the Laplace operator (Proc. Sympos. Pure Math., Univ. Hawaii, Honolulu, Hawaii, 1979)*, pages 205–239. Amer. Math. Soc., 1980.
- [69] Chen-Yun Lin, Arin Minasian, Xin Jessica Qi, and Hau-Tieng Wu. Manifold learning via the principle bundle approach. *Frontiers in Applied Mathematics and Statistics*, 4:21, 2018.
- [70] Chen-Yun Lin and Hau-Tieng Wu. Embeddings of riemannian manifolds with finite eigenvector fields of connection laplacian. *Calculus of Variations and Partial Differential Equations*, 57, 04 2018.
- [71] C.Y. Lin. Manifold embeddings by heat kernels of connection laplacian. <https://arxiv.org/submit/4064908/view>, 2022.
- [72] John Lott and Cédric Villani. Ricci curvature for metric-measure spaces via optimal transport. *Ann. of Math. (2)*, 169(3):903–991, 2009.
- [73] S Marchesini, Y.-C. Tu, and H.-T. Wu. Alternating Projection, Ptychographic Imaging and Phase Synchronization. *Appl. Comput. Harmon. Anal.*, 41(3):815–851, 2016.
- [74] H.P. McKean. An upper bound to the spectrum of δ on a manifold of negative curvature. *Journal of Differential Geometry*, 4:359–366, 1970.
- [75] Facundo Mémoli and Guillermo Sapiro. A theoretical and computational framework for isometry invariant recognition of point cloud data. *Found. Comput. Math.*, 5(3):313–347, 2005.
- [76] S Minakshisundaram. Zeta functions on the sphere. *J. Indian Math. Soc.*, 13:41–48, 1949.
- [77] Subbaramiah Minakshisundaram and Åke Pleijel. Some properties of the eigenfunctions of the laplace-operator on riemannian manifolds. *Canadian Journal of Mathematics*, 1(3):242–256, 1949.
- [78] Carlo Miranda. *Equazioni alle derivate parziali di tipo ellittico*. Ergebnisse der Mathematik und ihrer Grenzgebiete, (N.F.), Heft 2. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1955.

- [79] Peter Petersen. *Riemannian geometry*, volume 171 of *Graduate Texts in Mathematics*,. Springer, 2016.
- [80] Jacobus W. Portegies. Semicontinuity of eigenvalues under intrinsic flat convergence. *Calc. Var. Partial Differential Equations*, 54(2):1725–1766, 2015.
- [81] Jacobus W. Portegies. Embeddings of riemannian manifolds with heat kernels and eigenfunctions. *Communications on Pure and Applied Mathematics*, 69(3):478–518, 2016.
- [82] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500):2323–2326, 2000.
- [83] Richard M Schoen and Shing-Tung Yau. *Lectures on differential geometry*, volume 2. International press Cambridge, MA, 1994.
- [84] A. Singer. From graph to manifold Laplacian: the convergence rate. *Appl. Comput. Harmon. Anal.*, 21(1):128–134, 2006.
- [85] A Singer and H.-T. Wu. Vector Diffusion Maps and the Connection Laplacian. *Comm. Pure Appl. Math.*, 65(8):1067–1144, 2012.
- [86] A Singer and H.-T. Wu. Spectral Convergence of the Connection Laplacian from random samples. *Information and Inference: A Journal of the IMA*, 6(1):58–123, 2017.
- [87] Amit Singer and Hau-tieng Wu. Orientability and diffusion maps. *Appl. Comput. Harmon. Anal.*, 31(1):44–58, 2011.
- [88] Christina Sormani. *Perspectives in Scalar Curvature edited by Gromov and Lawson*, chapter Conjectures on Convergence and Scalar Curvature. World Scientific, 2022.
- [89] Christina Sormani and Stefan Wenger. The intrinsic flat distance between Riemannian manifolds and other integral current spaces. *J. Differential Geom.*, 87(1):117–199, 2011.
- [90] Karl-Theodore Sturm. On the geometry of metric measure spaces. I. *Acta Mathematica*, 196(1):133–177, 2006.
- [91] J. B. Tenenbaum, V. de Silva, and J. C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 290(5500):2319–2323, 2000.

- [92] S. R. S. Varadhan. Diffusion processes in a small time interval. *Communications on Pure and Applied Mathematics*, 20:659–685, 1967.
- [93] Sathamangalam R Srinivasa Varadhan. On the behavior of the fundamental solution of the heat equation with variable coefficients. *Communications on Pure and Applied Mathematics*, 20(2):431–455, 1967.
- [94] Cédric Villani. *Optimal transport. Old and new.*, volume 338 of *Grundlehren der mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2008.
- [95] Max Wardetzky, Saurabh Mathur, Felix Kälberer, and Eitan Grinspun. Discrete laplace operators: no free lunch. In *Symposium on Geometry processing*, pages 33–37. Aire-la-Ville, Switzerland, 2007.
- [96] Shing Tung Yau. Isoperimetric constants and the first eigenvalue of a compact Riemannian manifold. *Ann. Sci. École Norm. Sup. (4)*, 8(4):487–507, 1975.