# Explorations of Infinitesimal Inverse Spectral Geometry

by

Mikhail Panine

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Mathematics
in
Applied Mathematics

Waterloo, Ontario, Canada, 2013

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

# Abstract

Spectral geometry is a mathematical discipline that studies the relationship between the geometry of Riemannian manifolds and the spectra of natural differential operators defined on them. The spectra of Laplacians are the ones most studied in this context. A sub-field of this discipline, called inverse spectral geometry, studies how much geometric information one can recover from such spectra.

The motivation behind our study of inverse spectral geometry is a physical one. It has recently been proposed that inverse spectral geometry could be the missing mathematical link between quantum field theory and general relativity needed to unify those theories into a single theory of quantum gravity [47]. Unfortunately, this proposed link is not well understood. Most of the efforts in inverse spectral geometry were historically concentrated on the generation of counterexamples to the most general formulation of inverse spectral geometry and the few positive results that exist are quite limited. In order to remedy to that, it has been proposed to linearize the problem, and study an infinitesimal version of inverse spectral geometry [47].

In this thesis, I begin by reviewing the theory of pseudodifferential operators and using it to prove the spectral theorem for elliptic operators. I then define the commonly used Laplacians and survey positive and negative results in inverse spectral geometry. Afterwards, I briefly discuss a coordinate free reformulation of Riemannian geometry via the notion of spectral triple. Finally, I introduce a formulation of inverse spectral geometry adapted for numerical implementations and apply it to the inverse spectral geometry of a particular class of star-shaped domains in $\mathbb{R}^2$.

## Acknowledgements

# Table of Contents

# List of Figures

# Chapter 1

# Introduction

One of the most important unresolved problem of modern theoretical physics is the unification of general relativity and quantum field theory into a single, consistent, theory of quantum gravity. This problem has remained open for the past half-century despite the efforts of some of the best theoretical physicists. Part of the difficulty of the formulation of quantum gravity resides in the vast difference between the mathematical disciplines of differential geometry and functional analysis, the languages of general relativity and quantum theory, respectively.

It has been recently proposed by Kempf [47, 48] that bridging this vast mathematical gap is key to the long sought unification. His proposal is to use results from a mathematical discipline known as inverse spectral geometry (ISG). As famously put by Mark Kac [45], inverse spectral geometry is the quest to answer the question "Can one hear the shape of a drum?"[1] The drum in question is the studied manifold, and the sound of the drum are the spectra of some differential operators defined on it. Chief amongst them is the spectrum of the Laplace-Beltrami operator, a generalization of the familiar Laplacian. Other Laplacians, such as the Hodge Laplacian on $p$-forms are also used in the literature, but to a lesser extent. Spectral geometry thus naturally contains the differential geometric language of general relativity and the spectral, functional analytic, language of quantum theory.

Kac's original motivation for asking his question was the then known fact that the asymptotic properties of the spectrum of the Laplacian of a domain in $\mathbb{R}^N$ with smooth boundary can be used to determine the volumes of both the domain and the boundary.

---

[1] The author desperately tried to break with nearly 50 years of tradition by not citing Kac's famous paper on the first page. Manifestly, his best efforts were met with failure.

Those results are discussed in Chapter 3. For now, we set them aside and explore a different, quantum field theoretic, motivation due to Kempf [47, 48] for why it is plausible that the shape of a manifold could be determined from the spectra of Laplace operators.

Consider the surface of an arbitrarily shaped potato. In mathematical terms, one would usually describe this potatoid as a differentiable two dimensional manifold $\mathcal{M}$ with a Riemannian metric $g$. Now, let's imagine that one wants to record the shape of the potato by selecting $M$ generic points on its surface and measuring their pairwise distances. The distance between the points $i$ and $j$ is recorded as the entry $G_{ij}$ of an $M \times M$ matrix $G$. One expects that if $M$ is large enough and the points are chosen in a sufficiently even manner, the matrix $G_{ij}$ would allow one to reconstruct the shape of the potatoid up to small details not captured by $G$. This can be put differently using an observation due to Helmholtz. Pick $M$ points of $\mathbb{R}^N$ and record their coordinates $\{x_\mu^{(i)}\}_{\mu=1\ldots N}^{i=1\ldots M}$. By Pythagoras' theorem, the squares of the distances between the points are expressed as $M(M-1)/2$ equations depending upon the $MN$ coordinates.

$$G_{ij}^2 = \sum_{\mu=1}^{N} (x_\mu^{(i)} - x_\mu^{(j)})^2 \tag{1.1}$$

We want to pick sufficiently many points $M$ to eliminate the $MN$ coordinates in order to recast the shape strictly in terms of the mutual distances. We thus want the number of equations to be greater than the number of unknowns $MN$. Explicitly, $M(M-1)/2 \geq MN$ or, equivalently, $M > 2N + 1$. This leaves $M(M-1)/2 - MN$ equations that necessarily hold, by Pythagoras' law. If, however, the points were not picked in a flat manifold, the remaining equations will be violated. The way in which those equations fail to hold is a manifestation of the curvature of the manifold.

In order to recast the above discussion in terms of the spectrum of a Laplace operator, it is necessary to substitute the mutual distances used above with correlations of quantum fields. Consider a scalar quantum field $\hat{\phi}$. Its equal-time 2-point correlation function is given by

$$G(x, x') = \langle 0|\hat{\phi}(x)\hat{\phi}(x')|0\rangle \tag{1.2}$$

It is well known that it decays as the distance between $x$ and $x'$ increases [6]. One can thus use $G(x, x')$ as a measure of distance, at least to an extent. Suppose that the considered spacetime is a compact Riemannian manifold $(\mathcal{M}, g)$ without boundary, either in the sense of compact constant time slices or a Wick rotated Euclidean signature spacetime. The

compactness requirement ensures that Laplace operators on that manifold have discrete spectrum. In quantum field theoretic language, this is an infrared cutoff. Let $\{x_i\}_{i=1}^M$ be a set of uniformly chosen points in $\mathcal{M}$. By evaluating $G(x, x')$ at the chosen sample points one obtains $G_{ij} = G(x_i, x_j)$, an analogue to the mutual distance matrix used above. It is thus expected that this matrix can be used to reconstruct the shape of $(\mathcal{M}, g)$ up to a certain cutoff scale.

Of course, one could choose a different set of $M$ sample points $\{\tilde{x}_i\}_{i=1}^M$ and obtain a different matrix of correlations $\tilde{G}_{ij} = G(\tilde{x}_i, \tilde{x}_j)$. By the heuristic argument above, both $G_{ij}$ and $\tilde{G}_{ij}$ must describe the same manifold $(\mathcal{M}, g)$, up to some finite precision. The matrices $G_{ij}$ and $\tilde{G}_{ij}$ are thus bound to be somehow related. More specifically, part of the information they contain has to be common geometric information about the manifold, while the remainder would be a function of the choice of sampling points. This is entirely analogous to the usual situation in differential geometry where the metric tensor encodes both the shape of the manifold and the peculiarities of the chosen coordinate system. In the case of the matrices $G_{ij}$ and $\tilde{G}_{ij}$, it is possible to disentangle the two kinds of information encoded in the matrix of correlations by considering it in a functional analytic picture. Let $\hat{\phi}$ be massless. Recall that $G(x, x')$ is then related to the inverse of the Laplacian,

$$\Delta G(x, x') = \delta(x - x') \tag{1.3}$$

Alternatively, if $|x)$ denotes a position eigenvector, this is expressed as

$$G(x, x') = (x|\Delta^{-1}|x') \tag{1.4}$$

Let $\{\psi_i\}_{i=1}^M$ be the $M$ eigenvectors of $\Delta$ corresponding to the $M$ lowest eigenvalues and let $P$ be the projector onto the subspace spanned by those vectors. In quantum field theoretic language, this corresponds to a spatially covariant ultraviolet cutoff. Necessarily, $G(x, x')$ must be modified as

$$G(x, x') = (x|P\Delta^{-1}P|x') \tag{1.5}$$

The corresponding matrix of correlations is thus

$$G_{ij} = G(x_i, x_j) = (x_i|P\Delta^{-1}P|x'_j) \tag{1.6}$$

It is purposefully arranged that the number of sampling points coincides with the number of eigenvectors of the Laplacian left after the cutoff is imposed. Indeed, the components of

$G_{ij}$ can then be seen as the matrix elements of the operator $P\Delta^{-1}P$ in the basis $\{P|x_i\rangle\}_{i=1}^M$. Consequently, using a different set of sampling points corresponds to expressing $P\Delta^{-1}P$ in a different basis $\{P|\tilde{x}_i\rangle\}_{i=1}^M$. The geometric information contained in $G_{ij}$ is thus the information that is left invariant under change of basis. From basic linear algebra, it is known that the spectrum of $P\Delta^{-1}P$ is precisely that information. Since $\Delta^{-1}$ is the (pseudo)inverse of $\Delta$, the spectrum of $P\Delta^{-1}P$ corresponds to the first $M$ eigenvalues of $\Delta$. As $M$ can be taken arbitrarily high, Kempf then concludes that the spectrum of the Laplacian $\Delta$ can plausibly encode the shape of $(\mathcal{M}, g)$. Of course, this argument is only heuristic. It nonetheless provides a justification to why at least some geometric features of $(\mathcal{M}, g)$ can be recovered from the spectrum of $\Delta$.

Sadly, this program is not achievable in the current state of knowledge. First, there are known counterexamples to the most general possible formulation of inverse spectral geometry. Second, even in the cases when some identification of shape from spectra is known to be possible, restrictive symmetry conditions need to be imposed upon the set of considered shapes to make the problem tractable. In terms completely detached from the nature of spectral geometry, one could say that the nonlinearity of the map between shapes and spectra is to blame for all of its difficulties.

To remedy to this problem, Kempf [47] proposed to use a linearized, infinitesimal version of inverse spectral geometry, iterating small changes in shape to get desired small changes in spectrum. In that way, he conjectured, one could get from some starting shape to a target shape with a desired spectrum. This method has been numerically implemented and successfully applied to a class of two-dimensional manifolds discretized as graphs [1]. Our work goes in the same direction, as we numerically apply infinitesimal inverse spectral geometry to a class of domains in $\mathbb{R}^2$.

The structure of the present thesis is as follows. In Chapter 2, we prove the spectral theorem for elliptic operators on smooth vector bundles of a compact Riemannian manifold without boundary. This is a strictly technical step on the road to spectral geometry. Indeed, to study how spectra of operators reflect shape, it is necessary to first ascertain that such spectra exist. There are multiple paths to this result and we choose that of the theory of pseudodifferential operators. We then review the heat kernel and its asymptotic expansion, an important tool used to extract geometric information from the spectra of elliptic operators. In Chapter 3, we recall the definitions of various Laplacians and review the geometric information one can extract from their spectra via the asymptotic expansion of the heat kernel. Chapter 4 is dedicated to a review of inverse spectral geometry. There we survey important counterexamples and positive results in both inverse spectral geometry and some analogous problems, such as the determination of a Schrödinger potential from the spectrum of the quantum mechanical Hamiltonian. Chapter 5 stands apart from

the others, as it reviews the seemingly unrelated subject of spectral triples, an algebraic description of Riemannian manifolds. It is our belief that spectral triples could be a valuable mathematical tool in subsequent work on infinitesimal inverse spectral geometry. Finally in Chapter 6 we propose a tentative definition of infinitesimal inverse spectral geometry adapted for numerical computations. We then apply our definition to the inverse spectral geometry of a particular class of star-shaped domains in $\mathbb{R}^2$ using the spectrum of the Laplace-Beltrami operator with Dirichlet boundary conditions. Chapter 7 concludes our thesis by a discussion of future directions of research in inverse spectral geometry.

# Chapter 2

# Elliptic Operators and Heat Equations

In order to study spectral geometry, one must first study spectral theory. Indeed, in order to discuss the spectra of operators, one must first make sure that they exist. The goal of this chapter is to obtain spectral theorems valid for all Laplacians, regardless of the vector bundle upon sections of which they act. This can be achieved by studying elliptic partial differential operators on arbitrary smooth vector bundles, as all Laplacians are such operators and ellipticity turns out to be the key to their properties. Our exploration of the theory of elliptic operators will take us to the realm of pseudodifferential operators, a powerful generalization of both differential and integral operators. This is not the only possible path to the spectral theorem for Laplacians, as will be briefly discussed later.

The presentation closely follows [25] and [54]. Many additional comments are drawn from an alternative presentation found in [59], which never invokes the notion of pseudodifferential operator. We consider the proofs that are provided in this chapter to be of interest because they either illustrate the power of the theory of pseudodifferential operators or are particularly illuminating. Since the goal is to give the reader an overall impression of the theory of pseudodifferential operators, lemmas judged too technical are omitted.

### Notation

At this point it is necessary to introduce the multi-index notation. A multi-index $\alpha$ is an $N$-tuple $(\alpha_1, ..., \alpha_N)$ such that $\alpha_i \in \mathbb{N}$. We consider that $0 \in \mathbb{N}$. The norm of the multi-index $\alpha$ is taken to be $|\alpha| = \alpha_1 + \alpha_2 + ... + \alpha_N$. The factorial of the multi-index is

$\alpha! = \alpha_1! \cdots \alpha_N!$. For multi-indices $\alpha, \beta, \gamma$ such that $\alpha = \beta + \gamma$, one defines $\binom{\alpha}{\beta} = \alpha!/(\beta!\gamma!)$. The multi-index $\alpha$ and its norm are used to compactly denote powers of vectors in $\mathbb{R}^N$ and certain compositions of partial differential operators. Let $x = (x_1, ..., x_N)$ denote vectors in $\mathbb{R}^N$ and $\partial_i^x = \partial/\partial x_i$. We define the following shorthand expressions:

$$
\begin{aligned}
x^\alpha &= x_1^{\alpha_1} \cdots x_N^{\alpha_N} \\
d_x^\alpha &= (\partial_1^x)^{\alpha_1} \cdots (\partial_N^x)^{\alpha_N} \\
D_x^\alpha &= (-i)^{|\alpha|} d_x^\alpha
\end{aligned}
\tag{2.1}
$$

The second notation introduced for the composition of partial derivatives will simplify expressions involving Fourier transforms. For $f, g \in C^\infty(\mathbb{R}^N)$, the Leibnitz rule and Taylor's theorem take the following form:

$$
D_x^\alpha(fg) = \sum_{\beta+\gamma=\alpha} \binom{\alpha}{\beta} (D_x^\beta f)(D_x^\gamma g)
\tag{2.2}
$$

$$
f(x) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!} d_x^\alpha f(x_0)(x - x_0)^\alpha + \mathcal{O}(|x - x_0|^{n+1})
\tag{2.3}
$$

## 2.1 Sobolev Spaces

In this section, we define the function spaces needed to study the theory of pseudo-differential operators. We start by recalling the definitions and properties of Schwartz function spaces and Fourier transforms. Fourier transforms will be central to the definition of pseudo-differential operators, as the Fourier transform of partial derivatives of a function takes a particularly simple form. We then define Sobolev Spaces, a particular class of function spaces adapted for the study of differentiability of solutions of partial differential equations. The presentation follows [25, 54].

### 2.1.1 Fourier Transform

The definition of the Fourier transform requires the introduction of a few spaces of smooth functions. The space of smooth $\mathbb{C}$-valued functions of $\mathbb{R}^N$ is denoted $C^\infty(\mathbb{R}^N)$. Compactly supported smooth functions are denoted $C_0^\infty(\mathbb{R}^N)$. When specified, the same notation is used for real valued smooth functions.

**Definition 2.1.1** (Schwartz Class). *The Schwartz class $\mathcal{S}$ is the vector space of all smooth complex valued functions $f : \mathbb{R}^N \to \mathbb{C}$ such that for any multi-indices $\alpha, \beta$, there are constants $C_{\alpha,\beta}$ such that*

$$|x^\alpha D_x^\beta f| \leq C_{\alpha,\beta} \tag{2.4}$$

Notice that $C_0^\infty \subset \mathcal{S}$. A consequence of the above definition is that functions in $\mathcal{S}$ have derivatives that decrease at $\infty$ faster than the inverse of any polynomial. This can be expressed by the following estimate:

$$|D_x^\beta f| \leq C_{n,\beta}(1 + |x|)^n \quad, \quad \forall(n, \beta) \tag{2.5}$$

Following the clever suggestion in [25], we define an integral measure on $\mathbb{R}^N$ that simplifies the formulas related to Fourier transforms:

$$dF_x = (2\pi)^{-N/2} dx^N \tag{2.6}$$

In the above formula, $dx^N$ denotes the usual measure on $\mathbb{R}^{\mathbb{N}}$. The symbol $dF_x$ is not a standard notation and is to be read "Fourier transform adapted measure over the $x$ coordinates". For $f, g \in \mathcal{S}$ the $L_2$ inner product is defined as follows:

$$(f, g)_{L_2} = \int f(x)g^*(x)dF_x \tag{2.7}$$

This inner product is clearly equivalent to the standard one. The space of square-integrable functions $L_2(\mathbb{R}^N)$ can be defined as the completion of $\mathcal{S}$ with respect to this inner product. This space can also be defined as the completion of $C_0^\infty(\mathbb{R}^N)$ as this last space is dense in $\mathcal{S}$ with respect to the $L_2$ inner product.

**Definition 2.1.2** (Convolution). *For $x, y \in \mathbb{R}^N$ and $f, g \in \mathcal{S}$, the convolution product is defined as:*

$$(f * g)(x) = \int f(x - y)g(y)dF_y = \int f(y)g(x - y)dF_y \tag{2.8}$$

*It can be shown that $(f * g) \in \mathcal{S}$. This definition can be extended to $f, g \in L_2(\mathbb{R}^{\mathbb{N}})$ by density of $\mathcal{S}$ in $L_2(\mathbb{R}^{\mathbb{N}})$.*

**Definition 2.1.3** (Fourier Transform). *The Fourier transform of $f \in \mathcal{S}$ is defined as*

$$\hat{f}(\xi) = \mathcal{F}(f)(\xi) = \int e^{-ix \cdot \xi} f(x)dF_x \tag{2.9}$$

*where $x, \xi \in \mathbb{R}^N$. This can be extended to $f \in L_2(\mathbb{R}^{\mathbb{N}})$ by density of $\mathcal{S}$ in $L_2(\mathbb{R}^{\mathbb{N}})$.*

8

The Fourier transform enjoys a number of well-known useful properties:

**Lemma 2.1.1.** *The Fourier transform is a homeomorphism of $\mathcal{S}$ such that, for any mulit-index $\alpha$ and $f \in \mathcal{S}$:*

*(a) Fourier Inversion Formula:*

$$f(x) = \int \exp(ix \cdot \xi) \hat{f}(\xi) dF_\xi = \int \int e^{i(x-y) \cdot \xi} f(y) dF_y dF_\xi \qquad (2.10)$$

*(b) Interchange of differentiation and multiplication:*

$$D_\xi^\alpha \hat{f}(\xi) = \int (-1)^{|\alpha|} x^\alpha e^{-ix \cdot \xi} f(x) dF_x \qquad (2.11)$$

$$\xi^\alpha \hat{f}(\xi) = \int e^{-ix \cdot \xi} D_x^\alpha f(x) dF_x \qquad (2.12)$$

$$D_x^\alpha f(x) = \int e^{ix \cdot \xi} \xi^\alpha \hat{f}(\xi) dF_\xi \qquad (2.13)$$

$$x^\alpha f(x) = \int (-1)^{|\alpha|} e^{ix \cdot \xi} D_\xi^\alpha \hat{f}(\xi) dF_\xi \qquad (2.14)$$

*(c) Interchange of multiplication and convolution:*

$$\hat{f}\hat{g} = \mathcal{F}(f * g) \qquad (2.15)$$

$$\hat{f} * \hat{g} = \mathcal{F}(fg) \qquad (2.16)$$

*(d) Plancherel formula:*

$$(f, g)_{L_2} = (\hat{f}, \hat{g})_{L_2} \qquad (2.17)$$

*Proof.* See [25]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

## 2.1.2  Sobolev Spaces on $\mathbb{R}^N$

Sobolev spaces are normed function spaces whose norm depends not only upon the values that a function takes, but also upon the values of its derivatives. This enforces a particular kind of convergence that requires a certain degree of smoothness on the part of the functions. Consequently, those spaces are instrumental in establishing regularity results for solutions of elliptic partial differential equations, that is, proving that those solutions are smooth.

**Definition 2.1.4** (Sobolev Space). *The Sobolev Space $H_s(\mathbb{R}^N)$ is defined as the completion of the Schwartz class $\mathcal{S}$ (or $C_0^\infty$) with respect to the following norm:*

$$\|f\|_s^2 = \int (1 + |\xi|^2)^s |\hat{f}(\xi)|^2 dF_\xi \ , \quad f \in \mathcal{S} \ ; \ s \in \mathbb{R} \tag{2.18}$$

*The parameter $s$ is said to be the order of the Sobolev space.*

**Remark 2.1.1.** *The inclusion and density of $\mathcal{S}$ (or $C_0^\infty$) in $H_s(\mathbb{R}^N)$ for all $s$ allows for the comparison of Sobolev spaces of different order.*

**Remark 2.1.2.** *The above definition is not the most general possible definition of Sobolev spaces. Notice that $H_0(\mathbb{R}^N)$ can be identified with $L_2(\mathbb{R}^N)$. One can similarly define Sobolev spaces that reduce to any $L_p$ space in the case $s = 0$. For integer $s$, such a construction can be found in, for instance, [59]. Sobolev spaces based on $L_2$ are particularly useful as they are Hilbert spaces.*

The meaning of the norm in (2.18) can be clarified by exhibiting equivalent norms. Specifically, one can relate the parameter $s$ to the number of derivatives of $f$ one desires to consider. Recall that two norms $\| \cdot \|_1$ and $\| \cdot \|_2$ are equivalent if there are positive constants $c_1$ and $c_2$ such that $c_1 \|v\|_1 \leq \|v\|_2 \leq c_2 \|v\|_1$ for all vectors $v$ in the normed space in question. It is straightforward to show that there exist constants $c_1$ and $c_2$ such that the following inequality holds:

$$c_1 (1 + |\xi|^2)^s \leq (1 + |\xi|)^{2s} \leq c_2 (1 + |\xi|^2)^s \tag{2.19}$$

This can be used to establish that for $f \in H_s(\mathbb{R}^N)$ the following norm is equivalent to the one defined in (2.18).

$$\|f\|_{s,eq} = \int (1 + |\xi|)^{2s} |\hat{f}(\xi)|^2 dF_\xi \ , \quad f \in \mathcal{S} \ ; \ s \in \mathbb{R} \tag{2.20}$$

The benefits of using this norm need not be clear at this point. If $s = k \in \mathbb{N}$, a procedure similar to the one employed to prove the above equivalence can be employed to obtain a more enlightening equivalent norm. First, notice that there are positive $c_1, c_2$ such that the following inequality holds:

$$c_1 (1 + |\xi|^2)^k \leq \sum_{|\alpha| \leq k} |\xi^\alpha| \leq c_2 (1 + |\xi|^2)^k \tag{2.21}$$

Note that the key ingredient to obtain such inequalities is the study of the terms having the highest order in $|\xi|$. The highest order part of polynomial expressions will become a recurrent theme later on. The previous inequality can be used to obtain the following one:

$$c_1\|f\|_k^2 \leq \int \sum_{|\alpha|\leq k} |\xi^\alpha \hat{f}| dF_\xi \leq c_2\|f\|_k^2 \tag{2.22}$$

Lemma 2.1.1 allows one to interchange differentiation and integration to obtain the following equivalent norm:

$$\|f\|_{k,eq}^2 = \sum_{\alpha\leq k} \int \|D_x^\alpha f\|^2 dF_x \tag{2.23}$$

**Remark 2.1.3.** *Equation* (2.23) *makes perfect sense on $\mathcal{S}$, but requires one to take derivatives of a priori non-differentiable functions when applied to $H_s$. Yet, we claim that the norm defined in* (2.23) *is equivalent to the one used to define $H_s$. This can be resolved by carrying out the construction of Sobolev spaces in reverse order. In fact, the norm in Equation* (2.23) *is the conventional starting point for the definition of Sobolev spaces. A notion of differentiation on $L_2(\mathbb{R}^N)$ is provided by the weak derivative. This construction is carried out with great clarity in* [59].

Equation (2.23) shows that the order $s$ of the Sobolev space $H_s(\mathbb{R}^N)$ can sometimes be interpreted as the number of derivatives taken into account when determining the norm of a function. Of course, since $s \in \mathbb{R}$, this interpretation must be augmented by the inclusion of derivatives of fractional and negative orders. Fractional order derivatives are of no immediate interest to our endeavours. Negative order ones can be understood as dual spaces to the positive order ones and will actually play a crucial role in the definition of inverses of differential operators. Recall that the continuous dual of a Hilbert space is the space of bounded linear functionals on that space. By the Riesz representation theorem, there is a canonical isometry between a Hilbert space and its continuous dual [10]. It is thus permissible to identify a Hilbert space with its continuous dual. In the setting of $H_s$ spaces, the following alternative characterization of the continuous dual holds:

**Lemma 2.1.2** (Dual Sobolev Spaces)**.** *$H_{-s}$ can be identified with the continuous dual of $H_s$.*

*Proof.* See [25]. $\square$

Just as $C^k$ spaces, Sobolev spaces of different orders are included in each other. Consider $s, t \in \mathbb{R}$ such that $s > t$. This ensures that $(1+|\xi|^2)^t \leq (1+|\xi|^2)^s$ and thus that $\|f\|_t \leq \|f\|_s$ for $f \in \mathcal{S}$. Since $\mathcal{S} \subset H_p(\mathbb{R}^N)$ for all $p \in \mathbb{R}$, one can use the identity map of $\mathcal{S}$ to map $\mathcal{S} \subset H_s(\mathbb{R}^N)$ into $\mathcal{S} \subset H_t(\mathbb{R}^N)$. This can be extended to a norm non-increasing injection $H_s(\mathbb{R}^N) \to H_t(\mathbb{R}^N)$. That is, every element of $H_s(\mathbb{R}^N)$ has a unique counterpart in $H_t(\mathbb{R}^N)$. This can be thought of as an inclusion relation. Also, note that since this map is linear and bounded, it is continuous. Similarly, one may extend the map $D_x^\alpha : \mathcal{S} \to \mathcal{S}$ to a map between Sobolev spaces of different order.

**Lemma 2.1.3** (Extension of $D_x^\alpha$). $D_x^\alpha : \mathcal{S} \to \mathcal{S}$ extends to a continuous map from $H_s(\mathbb{R}^N)$ to $H_{s-|\alpha|}(\mathbb{R}^N)$.

*Proof.* Since $H_s(\mathbb{R}^N)$ is linear, it remains to show boundedness to establish continuity. There exists a positive constant $c$ such that $|\xi^\alpha|^2(1+|\xi|^2)^{s-|\alpha|} \leq c(1+|\xi|^2)^s$. Thus one has

$$\|D_x^\alpha f\|_{s-|\alpha|}^2 = \int |\xi^\alpha \hat{f}|^2 (1+|\xi|^2)^{s-|\alpha|} dF_\xi \leq c\|f\|_s^2 \tag{2.24}$$

This shows boundedness and concludes the proof. □

**Remark 2.1.4.** *The simplicity of the above proof is to be taken as a testament of the power of using the Fourier transform to define the norm on $H_s(\mathbb{R}^N)$. One can find a much longer proof that uses the derivative norm(equation (2.23)) in [59].*

Our study of pseudo-differential operators requires two more classical results from the theory of Sobolev spaces. Consider the completion of $\mathcal{S}$ with respect to the following norm:

$$\|f\|_{\infty,k} := \sup_{x \in \mathbb{R}^N} \sum_{|\alpha| \leq k} |D_x^\alpha f| \tag{2.25}$$

The function space thus obtained is a subset of $C^k(\mathbb{R}^N)$ [25].

**Theorem 2.1.1** (Sobolev). *Let $k \in \mathbb{N}$ and let $s > k + \frac{1}{2}N$. If $f \in H_s(\mathbb{R}^N)$, then $f \in C^k(\mathbb{R}^N)$ and $\|f\|_{\infty,k} \leq C\|f\|_s$ for some positive constant $C$.*

*Proof.* See [25]. □

**Theorem 2.1.2** (Rellich Compactness). *Let $K$ be a compact subset of $\mathbb{R}^N$ and let $\{f_n\}$ be a sequence in $C_0^\infty(K)$. If $\{f_n\}$ is bounded by a positive constant $C$ ($\|f_n\| \leq C$), there exists a subsequence $\{f_{n_i}\}$, which converges in $H_t$ for any $t < s$.*

*Proof.* See [25]. □

**Remark 2.1.5.** *From the Rellich Compactness Theorem it follows that the natural injection from $H_s(K)$ to $H_t(K)$ is compact when $K \subset \mathbb{R}^N$ is compact.*

### 2.1.3 Sobolev Spaces on Manifolds

Let $\mathcal{M}$ be an $N$-dimensional compact Riemannian manifold *without boundary* with metric tensor $g$. The associated volume form shall be denoted $dV_g$. $C^\infty(\mathcal{M})$ shall denote the space of smooth functions on $\mathcal{M}$. Let $\{U_i, h_i\}$ be an atlas of $\mathcal{M}$ where $\{U_i\}$ is a finite open cover of $\mathcal{M}$ and $\{h_i : U_i \to \mathbb{R}^N\}$ are the associated coordinate charts. Let $\{\phi_i\}$ be a partition of unity subordinate to the cover $\{U_i\}$. For $f, g \in C^\infty(\mathcal{M})$, the $L_2$ inner product on $\mathcal{M}$ is defined to be

$$(f, g)_{L_2(\mathcal{M})} = \int_{\mathcal{M}} f(x) g^*(x) dV_g \tag{2.26}$$

$L_2(\mathcal{M})$ is defined to be the completion of $C^\infty(\mathcal{M})$ in the norm induced by the above inner product. Anticipating the definition of Sobolev spaces on manifolds, this norm shall be denoted $\| \cdot \|_0^2$. The $L_2$ inner product defines a formal adjoint $P^*$ to any operator $P : C^\infty(\mathcal{M}) \to C^\infty(\mathcal{M})$ via the identity $(Pf, g) = (f, P^*g)$.

Sobolev spaces $H_s(\mathcal{M})$ are defined using partitions of unity and the properties of Sobolev spaces on domains in $\mathbb{R}^N$.

**Definition 2.1.5** (Sobolev Spaces on Manifolds). *Let $\mathcal{U} = (U_i, h_i, \phi_i)_{i=1}^\nu$ be an atlas of a compact Riemannian manifold without boundary $(\mathcal{M}, g)$. For $f \in C^\infty(\mathcal{M})$, define the following norm:*

$$\|f\|_{\mathcal{U},s}^2 := \sum_i \|h_{i*}(\phi_i f)\|_s \tag{2.27}$$

*The Sobolev space $H_s(\mathcal{M})$ is defined to be the completion of $C^\infty(\mathcal{M})$ in this norm.*

The norm in the above definition seems to depend upon the atlas $\mathcal{U}$. The following two lemmata shows that this is not the case.

**Lemma 2.1.4** (Atlas Independence). *Changing the atlas in the norm in definition 2.1.5 gives rise to an equivalent norm.*

*Proof.* See [25]. □

The Sobolev spaces on $\mathcal{M}$ enjoy the same properties as those on subsets of $\mathbb{R}^N$. This is summarized by the following lemma.

**Lemma 2.1.5.** *Let $s, t \in \mathbb{R}$ such that $s > t$. The following hold:*

*(a) The identity map on $C^\infty(\mathcal{M})$ extends to a norm non-increasing map from $H_s(\mathcal{M})$ to $H_t(\mathcal{M})$.*

*(b) $H_{-s}(\mathcal{M})$ is the continuous dual of $H_s(\mathcal{M})$.*

*Proof.* See [25]. □

### 2.1.4 Sobolev Spaces on Vector Bundles

The Sobolev spaces defined in the preceding section are not sufficient for our purposes as they only describe $\mathbb{C}$(or $\mathbb{R}$)-valued smooth functions, while we plan to describe operators acting on sections of vector bundles, such as the Hodge Laplacian. Fortunately, the extension of the notion of Fourier Transform and of Sobolev spaces is straightforward. The results of the preceding section will now be shown to hold for smooth vector bundles. Begin by recalling the definition of a smooth vector bundle [59]:

**Definition 2.1.6** (Vector Bundle). *A vector bundle is a quadruple $(E, \pi, \mathcal{M}, V)$ such that:*

*(a) $E, M$ are smooth manifolds*

*(b) $V$ is a vector space over the field $\mathbb{K} = \mathbb{R}, \mathbb{C}$*

*(c) $\pi : E \to \mathcal{M}$ is a surjective submersion. For $U \subset \mathcal{M}$, set $E|_U = \pi^{-1}(U)$.*

*(d) There exists a trivializing cover $\{U_i\}$ of $\mathcal{M}$, i.e. a cover such that for every $U_i$ there is a diffeomorphism*

$$\begin{aligned} \Psi_{U_i} &: E|_U \to U \times V \\ x &\mapsto (\pi(v), \Phi_p^{U_i}(v)) \end{aligned} \tag{2.28}$$

*where $\Phi_p^{U_i}$ is a diffeomorphism $E_p \to V$ for any $p \in U_i$. Moreover, for two trivializing neighborhoods $U_i, U_j$ with nonempty intersection, the map $\Phi_{U_i U_j}(p) = \Phi_p^{U_i}(\Phi_p^{U_j})^{-1} : V \to V$ is a linear automorphism and the map $p \mapsto \Phi_{U_i U_j}(p) \in Aut(F)$ is smooth.*

The space of smooth sections (or vector fields) of a bundle $E$ will be denoted $C^\infty(E)$. The key message to retain is that complex vector bundles locally look like $\mathbb{C}^p$. This is exploited in [54] to enable the Fourier transform to act on vector bundles. We follow that discussion to define the *good presentation* of a vector bundle on a compact Riemannian manifold.

Let $(\mathcal{M}, g)$ be a compact Riemannian manifold equipped with a vector bundle $E$ of rank $p$. Suppose that $\mathcal{M}$ has a finite cover $\{U_i\}_{i=1}^q$ by *closed* coordinate unit balls. The coordinates on those balls will be denoted by $y$. That is, the charts are such that $y_i : U_i \to \bar{B}_N = \{y \in \mathbb{R}^N : |y| \le 1\}$. Now assume that this covering has too much overlap and that one could use a cover by open balls of radius $1/\sqrt{2}$ instead. That is, let $B_i = \{p \in U_i : |y_i(p)|^2 \le 1/2\}$ and assume that $\mathcal{M} = \cup_i^q B_i$. Notice that $U_i$ and $B_i$ share coordinates on their intersection. On each $U_i$ a smooth trivialization of $E$ (that is actually defined on an open set containing $U_i$) is then picked. This trivialization is consequently also valid on the corresponding $B_i$. Now, change coordinates on the $U_i, B_i$ pairs in the following way:

$$x_i = \frac{1}{\sqrt{1 - |y_i|^2}} y_i \tag{2.29}$$

This expands the coordinate systems of the $U_i$ to cover all of $\mathbb{R}^N$, while the coordinate systems of the $B_i$ are now open balls of radius 1. A smooth section of $E$ represented in the trivialization chosen above is now a bounded smooth function $\mathbb{R}^N \to \mathbb{C}^p$. Choosing a partition of unity $\{\chi_i\}_{i=1}^q$ subordinate to the cover $\{B_i\}_{i=1}^q$, any smooth section $u$ of $E$ can be written as $u = \sum_{i=1}^q \chi_i u$. Each term $\chi_i u$ is smooth and compactly supported within its $B_i$. Such a choice of two covers with matching coordinates and of a partition of unity is called a *good presentation* of $E$. When multiple vector bundles are studied simultaneously, they are taken to be in the same good presentation.

The usefulness of good presentations is that the study of sections of $E$ is reduced to that of compactly supported $\mathbb{C}^p$-valued functions of $\mathbb{R}^N$. The Fourier transform of such a function is simply taken to be the component-wise Fourier transform. Similarly, the Schwartz space of $\mathbb{C}^p$-valued functions is taken to be the space of component-wise Schwartz functions. All the results of the preceding sections now generalize to this setting, provided that matrix-valued functions are multiplied in the appropriate order.

## 2.2 Pseudo-Differential Operators

Having established the the properties of Sobolev spaces, we are in position to define the action of a particular class of linear operators upon those spaces. The operators in question

are known as pseudo-differential operators, or $\Psi$DOs, and they generalize the notion of differential operator to fractional and negative orders. Fractional orders of differentiation are of no interest to our endevours, so the results of this section will often be restricted to integer orders. Negative orders are however of great interest, as they allow to construct pseudoinverses or *parametrices* of certain pseudo-differential operators. The theory of pseudo-differential operators thus puts differential operators and their inverses on equal footing. Our discussion closely follows [54], with some elements taken from [25]. A reader interested in results unrestricted to integer orders is invited to consult [25].

## 2.2.1   Differential Operators and Symbols

Let $(\mathcal{M}, g)$ be a Riemannian manifold and let $E, F$ denote smooth vector bundles over $\mathcal{M}$ of ranks $p$ and $q$, respectively. Assume that every point $x \in \mathcal{M}$ has a neighborhood $U$ with coordinates $(x_1, ..., x_N)$ over which $E$ and $F$ have local trivializations $E|_U \to U \times \mathbb{C}^p$ and $F|_U \to U \times \mathbb{C}^q$.

**Definition 2.2.1** (Differential Operators). *Let $P : C^\infty(E) \to C^\infty(F)$ be a linear operator. $P$ is a differential operator of order $m$ if, on all trivializing neighborhoods $U$ as above, $P$ can be written in the form*

$$P = \sum_{|\alpha| \leq m} A^\alpha D_x^\alpha \tag{2.30}$$

*where $A^\alpha$ is a $q \times p$ matrix of smooth $\mathbb{C}$-valued functions and $A^\alpha \neq 0$ for some $\alpha$ with $|\alpha| = m$.*

**Remark 2.2.1** (Algebraic PDOs). *One can also provide an algebraic definition of differential operators on vector bundles. Notice that $C^\infty(E)$ and $C^\infty(F)$ have more structure than just that of a vector space over $\mathbb{C}$. Indeed, they also are modules over $C^\infty(\mathcal{M})$. Let $\mathbf{Op}(E, F)$ denote the space of $\mathbb{C}$-linear operators $C^\infty(E) \to C^\infty(F)$ and let $\mathbf{Hom}(E, F)$ be the space of $C^\infty(\mathcal{M})$-homomorphisms between $E$ and $F$. $\mathbf{Hom}(E, F)$ is clearly a subspace of $\mathbf{Op}(E, F)$. Define the space of zeroth order partial differential operators to be:*

$$\mathbf{PDO}^0(E, F) = \mathbf{Hom}(E, F) \tag{2.31}$$

*Notice that operators in $\mathbf{PDO}^0(E, F)$ commute with multiplication by $C^\infty(\mathcal{M})$. One inductively defines the space of partial differential operators of order $m$ and lower as*

$$\mathbf{PDO}^{(m)}(E, F) = \{T \in \mathbf{Op}(E, F) : [T, f] \in \mathbf{PDO}^{m-1}(E, F) \; \forall f \in C^\infty(\mathcal{M})\} \tag{2.32}$$

*The interested reader may find a detailed account of this approach to elliptic equations on manifolds in [59]. We revisit this approach in Remark 2.2.2.*

A local change of the trivializations $E|_U$ and $F|_U$ is expressed by smooth maps $g_E : U \to GL_p(\mathbb{C})$ and $g_F : U \to GL_q(\mathbb{C})$, where $GL$ is the general linear group. Under such a trivialization, $P$ has the form

$$P = g_F \left( \sum_{|\alpha| \leq m} A^\alpha D_x^\alpha \right) g_E^{-1} = \sum_{|\alpha| \leq m} \mathcal{A}^\alpha D_x^\alpha \tag{2.33}$$

where the $\mathcal{A}^\alpha$ are $p \times q$ matrices of smooth functions, as before. Consider the $\mathcal{A}^\alpha$ such that $|\alpha| = m$. Those matrices can only come from those terms that have not used up any orders of differentiation on $g_E^{-1}$, otherwise they would not correspond to $|\alpha| = m$. Consequently,

$$\mathcal{A}^\alpha = g_F A^\alpha g_E^{-1} \quad \text{for } |\alpha| = m \tag{2.34}$$

Now, consider a change of coordinates $\tilde{x} = \tilde{x}(x)$ on $U$. Using the chain rule for partial derivatives, the following expression for $P$ is obtained:

$$\begin{aligned} P &= \sum_{|\alpha| \leq m} A^\alpha(x(\tilde{x}))(-i)^{|\alpha|} \left[ \left( \sum_{k=1}^N \frac{\partial \tilde{x}_k}{\partial x_1} \partial_k^{\tilde{x}} \right)^{\alpha_1} \cdots \left( \sum_{k=1}^N \frac{\partial \tilde{x}_k}{\partial x_m} \partial_k^{\tilde{x}} \right)^{\alpha_m} \right] \\ &= \sum_{|\alpha| \leq m} \tilde{A}^\alpha D_{\tilde{x}}^\alpha \end{aligned} \tag{2.35}$$

Once again, consider the highest order matrices $\tilde{A}^\alpha$, $\alpha = m$. In order for them to preserve the highest available order of differentiation, the Jacobian factors in equation (2.35) must all remain undifferentiated. Consequently,

$$\tilde{A}^\alpha = \sum_{|\beta|=m} A^\beta \left[ \partial_x^\beta \tilde{x} \right]^\alpha \quad \text{for } |\alpha| = m \tag{2.36}$$

The factor $\left[ \partial_x^\beta \tilde{x} \right]^\alpha$ is the symmetrization of the $m^{\text{th}}$ tensor power of the Jacobian of the coordinate transformation. Symmetrization is manifest as multiplication by scalar functions commutes. Let $\odot$ denote the symmetric tensor product. Together, equations (2.34) and (2.36) show that the coefficients $\{A^\alpha\}_{|\alpha|=m}$ are a well defined section of the bundle $(\odot^m T\mathcal{M} \otimes \mathbf{Hom}(E, F))$.

17

**Definition 2.2.2** (Principal Symbol). *Let $P$ be a differential operator of order $m$ and let $\sigma(P) \in C^\infty((\odot^m T\mathcal{M} \otimes \mathbf{Hom}(E, F))$ be the section defined above. $\sigma(P)$ is called the principal symbol of $P$.*

Since $T^*\mathcal{M}$ is canonically isomorphic to $T\mathcal{M}$, the principal symbol may be recast as a section of $(\odot^m T^*\mathcal{M} \otimes \mathbf{Hom}(E, F))$. In turn, this can be recast as a homogeneous polynomial in the tangent space coordinate $\xi$ using the following polarization trick [59].

**Lemma 2.2.1** (Polarization Trick). *Let $V$ be a vector space over $\mathbb{C}$ and let $V^*$ be is dual. Let $\phi$ be a symmetric $k$ multilinear map*

$$\phi : V \times ... \times V \;\; _{k \; times} \to \mathbb{C} \tag{2.37}$$

*The map $\phi$ is uniquely determined by the following polynomial in $v \in V$:*

$$\mathcal{P}_\phi(v) = \phi(v, ...v) \tag{2.38}$$

*Proof.* Let $\{t_i\}_{i=1}^k$ be auxiliary real variables. The lemma follows from the following *polarization formula*:

$$\phi(v_1, ..., v_k) = \frac{1}{k!} \frac{\partial^k}{\partial_{t_1} ... \partial_{t_k}} \mathcal{P}_\phi(t_1 v_1 + ... + t_k v_k) \tag{2.39}$$

$\square$

Lemma 2.2.1 allows one to write a polynomial $\sigma_\xi(P)$ equivalent to the principal symbol. At a point $x \in \mathcal{M}$, given $\xi \in T_x^*\mathcal{M}$, the polynomial symbol gives a map $\sigma_\xi(P) : E_x \to F_x$. In a trivializing neighbourhood, as above, this polynomial takes the form

$$\sigma_\xi(P) = \sum_{|\alpha|=m} A^\alpha(x) \xi^\alpha \tag{2.40}$$

In the following, whenever we refer to a principal symbol, it will be in its polynomial form.

Principal symbols have a very nice behaviour under linear combination and composition of differential operators, as expressed by the next lemma.

**Lemma 2.2.2.** *Let $P : C^\infty(E) \to C^\infty(F)$, $P' : C^\infty(E) \to C^\infty(F)$ and $Q : C^\infty(F) \to C^\infty(G)$ be differential operators on smooth vector bundles $E, F, G$. Assume that $P$ and $P'$ are of the same order $m$. Then, for all $\xi \in T^*\mathcal{M}$ and for all $t, t' \in \mathbb{C}$, the following holds*

$$\sigma_\xi(tP + t'P') = t\sigma_\xi(P) + t'\sigma_\xi(P') \tag{2.41}$$
$$\sigma_\xi(Q \circ P) = \sigma_\xi(Q) \circ \sigma_\xi(P) \tag{2.42}$$

18

*Sketch of Proof.* The lemma follows from considerations similar to those that were used to establish that the principal symbol is a section of $(\odot^m T^*\mathcal{M} \otimes \mathbf{Hom}(E, F))$. $\square$

**Remark 2.2.2** (Algebraic Symbol)**.** *The above construction of the principal symbol may seem artificial, especially when it is taken from being a section of $(\odot^m T\mathcal{M} \otimes \mathbf{Hom}(E, F))$ to a section of $(\odot^m T^*\mathcal{M} \otimes \mathbf{Hom}(E, F))$. Once again, the equivalent algebraic construction proves to be enlightening. The notation used here is the same as in Remark 2.2.1. For $P \in \mathbf{PDO}^{(m)}(E, F)$ and $f \in C^\infty(\mathcal{M})$, let $ad(f)P = [P, f]$. Set $x \in \mathcal{M}$. For any $m$ functions $\{f_i\}_{i=1}^m \subset C^\infty(\mathcal{M})$, consider*

$$\left\{\frac{1}{m!}ad(f_1)...ad(f_m)P\right\}\bigg|_x : E_x \to F_x \tag{2.43}$$

*This expression is symmetric in the $f_i$ as $ad(f) \circ ad(g) = ad(g) \circ ad(f)$. Moreover, it can be shown [59] that it depends only upon $\{\xi_i := df_i(x)\}_{i=1}^m \subset T_x^*\mathcal{M}$, rather than upon the values of the $f_i$ at $x$. Consequently, this expression induces a linear map $\sigma(P)(\xi_1, ..., \xi_m) : E_x \to F_x$. By the polarization trick of Lemma 2.2.1, the study of this linear map is reduced to the study of the following polynomial*

$$\sigma_\xi(P) := \sigma(P)(\xi, ..., \xi) \tag{2.44}$$

*This polynomial is again called the principal symbol of $P$. It differs from the definition established before by a factor of $(-i)^m$. This is due to the fact that Definition 2.2.1 uses $D_x^\alpha$ rather than $d_x^\alpha$. This choice simplifies the application of Fourier transform techniques in the definition of pseudo-differential operators.*

## 2.2.2 Pseudo-Differential Operators

The goal of this section is to define pseudo-differential operators ($\Psi$DOs) and to establish their main properties. Key to this endeavour is the notion of full symbol of a differential operator, which expresses the way a pseudo-differential operator acts on the Fourier transform of a function. Following [25] and [54], we first motivate the definition of the full symbol for differential operators and then use the Fourier transform to generalize the notion. This is similar in spirit to the way in which fractional derivatives made their way into the definition of a Sobolev space in Definition 2.1.4. The study of operators can then be reduced to the study of their symbols.

Let $P : C^\infty(E) \to C^\infty(F)$ be a differential operator of order $m$ acting on sections smooth vector bundles $E$ and $F$ on $\mathbb{R}^N$. Let $u \in C^\infty(E)$ be Schwartz and let $\hat{u}$ be its Fourier transform. In a good presentation, let $P$ be expressed as in Definition 2.2.1 .

**Definition 2.2.3** (Total Symbol). *In a good presentation of the vector bundles $E, F$ over $\mathbb{R}^N$, the* total symbol *of a differential operator $P = \sum_{|\alpha| \leq m} A^\alpha D_x^\alpha$ is defined to be*

$$p(x, \xi) = \sum_{|\alpha| \leq m} A^\alpha \xi^\alpha \tag{2.45}$$

*It is obtained by formally replacing all occurrences of $D_x^\alpha$ with $\xi^\alpha$.*

Observe that by Lemma 2.1.1, the action of $P$ on $u$ can be expressed as

$$Pu(x) = \int e^{ix \cdot \xi} p(x, \xi) \hat{u}(\xi) dF_\xi \tag{2.46}$$

Thus, the symbol $p(x, \xi)$ expresses the action of $P$ on the Fourier transform of smooth sections of $E$. By restricting the sum in Definition to $|\alpha| = m$, one recovers the definition of the principal symbol of $P$.

Conceptually, equation (2.46) is all that is necessary to define pseudo-differential operators. It suffices to reverse the order in which objects are defined by putting symbols first and operators second.

**Definition 2.2.4** (Pseudo-differential Operators). *For $m \in \mathbb{R}$, let $S^m$ be the set of all symbols $p(x, \xi)$ such that:*

*(a) $p(x, \xi)$ is a smooth matrix valued function on $\mathbb{R}^N \times \mathbb{R}^N$.*

*(b) For all pairs of multi-indices $(\alpha, \beta)$, there exist constants $C_{\alpha, \beta}$ such that*

$$\left| D_x^\alpha D_\xi^\beta p(x, \xi) \right| \leq C_{\alpha, \beta} (1 + |\xi|)^{m - |\beta|} \tag{2.47}$$

*Let $p \in S^m$. In a good presentation of $E, F$, the associated operator $P : C^\infty(E) \to C^\infty(F)$ is defined to be:*

$$Pf(x) := \int e^{ix \cdot \xi} p(x, \xi) \hat{f}(\xi) dF_\xi \tag{2.48}$$

*Such operators are said to be pseudo-differential operators of order $m$. The space of these operators is denoted $\Psi^m$. For two given vector bundles $E$ and $F$, $S = \cup_m S^m$ and $\Psi = \cup_m \Psi^m$ are the spaces of all symbols and all pseudo-differential operators, respectively. Conventionally, pseudo-differential operators are written as majuscules $(P, Q, ...)$, while their symbols are denoted by the corresponding minuscules $(p, q, ...)$.*

**Remark 2.2.3.** *The definition of the principal symbol for a pseudo-differential operator is more subtle than that of the principal symbol of a differential operator. It will have to wait.*

**Remark 2.2.4.** *Note that $m$ is not required to be a natural number. This allows operators that can be seen as fractional or negative partial derivatives to be defined.*

In the same fashion as the derivative operators of Lemma 2.1.3, the action of pseudo-differential operators extends to Sobolev spaces.

**Lemma 2.2.3** (Extension to $H_s$)**.** *Let $p \in S^m$ and $P \in \Psi^m$ be the associated operator. For any $s \in \mathbb{R}$, there exists a constant $C$ such that $\|Pf\|_{s-m} \leq C\|f\|_s$ for all $f \in \mathcal{S}$. $P$ then extends to a continuous linear map from $H_s$ to $H_{s-m}$.*

*Proof.* See [25] or [54]. □

Observe that by lemma 2.2.3, $\Psi$DOs of negative order map functions into Sobolev spaces of higher order. Thus, they add orders of differentiability to functions. Consequently, pseudo-differential operators of negative orders are known as smoothing operators. Consider the intersection of the spaces of symbols and of pseudo-differential operators of all orders.

$$S^{-\infty} = \bigcap_m S^m \quad , \quad \Psi^{-\infty} = \bigcap_m \Psi^m \tag{2.49}$$

Operators in $\Psi^{-\infty}$ are infinitely smoothing, that is they send elements of Sobolev spaces to smooth functions. This can be seen in the following way. Fix any $s \in \mathbb{R}$ and let $P \in \Psi^{-\infty}$. By definition of $\Psi^{-\infty}$, $P$ is in particular in $\Psi^m$ for some $m \in \mathbb{R}$. By lemma 2.2.3, $P$ can be extended to a continuous linear map between $H_s$ and $H_{s-m}$. Since $s$ and $m$ are both arbitrary real numbers, $P$ can really be extended to a continuous linear map between $H_s$ and $H_t$ for any $s, t \in \mathbb{R}$. Choose $k \in \mathbb{N}$ and let $t > k + \frac{1}{2}N$. By theorem 2.1.1, this implies that $Pf \in C^k$ for any $f \in H_s$. Since $k$ is arbitrary, this is valid for all $k \in \mathbb{N}$ and one has that $P(H_s) \subseteq C^\infty(E)$. $P$ is indeed infinitely smoothing. The following property of infinitely smoothing operators will be useful when discussing elliptic operators.

**Lemma 2.2.4** (Compactness of $\Psi^\infty$)**.** *Let $P : C^\infty(E) \to C^\infty(E)$ be an infinitely smoothing operator. Then, its extension onto Sobolev spaces $P : H_s(E) \to H_t(E)$ is compact for any $s, t$.*

*Proof.* By Lemma 2.2.3, $P$ can be extended to a bounded linear map $P : H_s(E) \to H_t(E)$ for any $s, t$. Since $P$ is infinitely smoothing, its image is in all Sobolev spaces, irrespective of the domain. Thus, one can temporarily consider that $P : H_s(E) \to H_{t+1}(E)$ and then use extension of the identity map of $C^\infty(E)$ to project $H_{t+1}(E)$ onto $H_t(E)$. By Theorem 2.1.2, this is a compact injection. Thus, $P$ is also compact. $\square$

In order to fully replace the study of operators with the study of symbols, it is necessary to express composition of operators and the operation of taking the adjoint of an operator in terms of symbols. Moreover, it is necessary to obtain the law of transformation of full symbols under coordinate change. Except for the principal symbol of a differential operator, this is no simple task. Fortunately, it turns out that precise determination of such expressions is not necessary. Instead, expressions will only be obtained up to the following equivalence relation.

**Definition 2.2.5** (Equivalence of Symbols)**.** *Let $p, q \in S$. If $p - q \in S^{-\infty}$, it is said that $p \sim q$.*

Because of this equivalence relation, operators in $\Psi^{-\infty}$ are sometimes known as *negligible* [32]. This equivalence relation allows one to expand the symbol of a pseudo-differential operator in a formal series.

**Definition 2.2.6** (Formal Development)**.** *Let $\{p_j\}$ be a sequence of symbols such that $p_j \in S^{m_j}$ and $m_j \downarrow -\infty$ as $j \uparrow \infty$. For a symbol $p \in S$ we write $p \sim \sum_j p_j$ if for every $m$ there exists an integer $k(m)$ such that $k \geq k(m)$ implies*

$$p - \sum_{j \leq k} p_j \in S^m \tag{2.50}$$

**Remark 2.2.5.** *It is important to note that the sum $\sum_{j \leq k} p_j$ is not necessarily convergent. $p \sim \sum_j p_j$ only means that the difference between $p$ and a partial sum with sufficiently many term of the sequence $\{p_j\}$ is arbitrarily smoothing.*

**Lemma 2.2.5** (Completeness)**.** *Let $\sum_{j=1}^\infty p_j$ be a formal sum of symbols such that $p_j \in S^{m_j}$ and $m_j \downarrow -\infty$ as $j \uparrow \infty$. There exists a pseudo-differential operator $P$ such that this sum is its formal development. $P$ is unique up to equivalence and its symbol $p$ is in $S^{m_1}$.*

*Sketch of Proof.* This lemma can be proven by explicitly constructing the symbol of an operator from the terms in the sum. Let $\phi : \mathbb{R}^+ \to [0,1]$ be a smooth function such that $\phi(t) = 0$ for $t \leq 1$ and $\phi(t) = 1$ for $t \geq 2$. Let $\{r_i\}_{i=1}^{\infty}$ be a set of radii such that $r_j \to \infty$ as $j \to \infty$. Consider the following series:

$$p(x,\xi) = \sum_{j=1}^{\infty} \phi(|\xi|/r_j) p_j(x,\xi) \tag{2.51}$$

The selection function $\phi(|\xi|/r_j)$ ensures the pointwise convergence of the series, as for every point $(x,\xi)$ only a finite number of symbols $p_j(x,\xi)$ contribute to the sum. The remainder of the proof can be found in either [54] or [25]. □

The last lemma ensures that one can replace the task of obtaining a pseudo-differential operator with some desired properties with the simpler task of obtaining the formal development of the symbol of that operator, as long as one only seeks the operator up to equivalence. Moreover, the proof of the lemma provides an explicit way to construct the symbol of an operator in that equivalence class from the formal development. However, none of the above provides a way to obtain the formal development in the first place. This problem is solved by a clever, but technical theorem that states that a particular class of integral operators with smooth kernel are pseudo-differential operators whose symbols have formal expansions given in terms of derivatives of the kernels of those operators. Calculations of formal expansions then reduce to showing that an operator can be represented in the form of an integral operator with a kernel that satisfies the hypotheses of the theorem. For this reason, [54] calls it the *Workhorse Theorem*.

The Workhorse Theorem can be used to prove a score of fundamental results. First come results about the localization of pseudo-differential operators.

**Definition 2.2.7** ($\varepsilon$-locality). *Let $A \subset \mathbb{R}^N$ be open and let $\varepsilon > 0$. Define $A_\varepsilon = \{x \in \mathbb{R}^N : d(x,A) \leq \varepsilon\}$. An operator $P : \mathcal{S} \to \mathcal{S}$ is said to be $\varepsilon$-local if for all $u \in C_0^{\infty}$,*

$$supp(Pu) \subset (supp(u))_\varepsilon \tag{2.52}$$

**Lemma 2.2.6** ($\varepsilon$-locality of $\Psi$DOs). *Let $P \in \Psi^m$ be such that its symbol $p(x,\xi)$ has compact $x$ support. Given any $\epsilon > 0$, there exists an $\varepsilon$-local operator $P_\varepsilon \in \Psi^m$ such that $P \sim P_\varepsilon$*

*Proof.* Uses the Workhorse Theorem. See [54]. □

**Lemma 2.2.7** (Localization of $\Psi$DOs). *Let $\chi = (\chi_1, \chi_2)$ be a pair of real-valued smooth functions on $\mathbb{R}^N$ with intersecting compact support. Let $P \in \Psi^m$. Define $P^\chi u = \chi_1 P(\chi_2 u)$. $P^\chi$ is also in $\Psi^m$. $P^\chi$ is sometimes referred to as a localization of $P$ [25].*

*Proof.* Uses the Workhorse Theorem. See [54]. $\square$

The Workhorse Theorem also yields results on the composition and adjoints of pseudo-differential operators. In order to properly discuss composition, it is necessary to restrict the spatial domains of operators. The set of all symbols of order $m$ that have compact $x$ support in a compact set $K \subset \mathbb{R}^N$ is denoted $S^m(K)$. The corresponding space of operators is denoted $\Psi^m(K)$. The notation for spaces of all orders and infinitely smoothing operators generalizes in a natural way:

$$S(K) = \bigcup_m S^m(K) \quad, \quad \Psi(K) = \bigcup_m \Psi^m(K)$$
$$S^{-\infty}(K) = \bigcap_m S^m(K) \quad, \quad \Psi^{-\infty}(K) = \bigcap_m \Psi^m(K)$$

(2.53)

**Lemma 2.2.8** (Symbol of a Composition). *Let $P \in \Psi^l(K)$ and $Q \in \Psi^m(K)$ have symbols $p$ and $q$, respectively. The symbol $s$ of $S = P \circ Q \in \Psi^{l+m}(K)$ has the following formal development:*

$$s \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha p)(D_x^\alpha q)$$

(2.54)

*Proof.* Uses the Workhorse Theorem. See [54]. $\square$

**Lemma 2.2.9** (Symbol of the Adjoint). *Let $P \in \Psi^m(K)$ have symbol $p$. Define $P^*$ to be the formal $L_2$ adjoint of $P$. That is, for all $u, v \in \mathcal{S}$ with compact support in $K$,*

$$(Pu, v)_{L_2(F)} = (u, P^* v)_{L_2(E)}$$

(2.55)

*The symbol $p^*$ of $P^*$ has the following formal development:*

$$p^* \sim \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} D_\xi^\alpha D_x^\alpha p^\dagger$$

(2.56)

*where $p^\dagger$ denotes the conjugate transpose of $p$.*

24

*Proof.* Uses the Workhorse Theorem. See [54]. □

The Workhorse Theorem is also used to establish the properties of pseudo-differential operators and their symbols under coordinate change.

**Lemma 2.2.10** (Coordinate Change). *Let $\phi : U \to V$ be a diffeomorphism between two open subsets of $\mathbb{R}^N$. Let $\phi^*$ and $\phi_*$ be the function (section) pull-back and push-forward, respectively. For each compact $K \subset U$, $\phi$ induces the following operator push-forward map:*

$$
\begin{aligned}
\phi_* &: \Psi^m(K) \to \Psi^m(\phi(K)) \\
P &\mapsto \phi_* \circ P \circ \phi^*
\end{aligned}
\tag{2.57}
$$

*Proof.* Uses the Workhorse Theorem. See [54]. □

The above lemma shows that coordinate changes in operators do not take a particularly complicated form. Symbols, however, change in most confusing ways. In a sense, this is not different from how symbols of differential operators transform under diffeomorphism. Recall that only the principal symbol had a tractable transformation law, which was not hard to establish from chain rule considerations. An analogous result exists for pseudo-differential operators. It however is much less straightforward to obtain, as one no longer has clear-cut orders of differentiation and the Leibniz rule needs not to apply. Moreover, symbol calculus has pushed all calculations into the realm of asymptotic expansions. Once again, the Workhorse Theorem comes to the rescue by helping to obtain a notion of *principal symbol* of a pseudo-differential operator.

**Definition 2.2.8** (Principal Symbol). *Let $P \in \Psi^m$ have symbol $p \in S^m$. The principal symbol of $P$ is the equivalence class $\sigma_\xi(P) = [p] \in S^m/S^{m-1}$.*

**Lemma 2.2.11.** *The principal symbol $\sigma_\xi(P)$ changes as a function of the cotangent bundle of $\mathbb{R}^N$ under diffeomorphisms. Let $(\tilde{x}, \zeta)$ denote the new cotangent space coordinates and let $(\partial x/\partial \tilde{x})^t$ be a transposed Jacobian matrix. Explicitly,*

$$
sym(\phi_*(P)) = p\left( x(\tilde{x}), \left(\frac{\partial x}{\partial \tilde{x}}\right)^t \zeta \right) \ mod(S^{m-1})
\tag{2.58}
$$

*If the diffeomorphism is linear, the above transformation law is exact.*

*Proof.* Uses the Workhorse Theorem. See [54]. □

**Remark 2.2.6.** *The above situation is entirely analogous to the one for differential operators. Intuitively, it corresponds to extracting a differential operator of order 1 out of a pseudo-differential operator of order $m$. The remainder is of course of order $m - 1$. Due to the fact that $m$ can be any real number, one is effectively allowed to conjure differential operators where there seemed to be none.*

Having established the properties of pseudo-differential operators under diffeomorphism on compact sets of $\mathbb{R}^N$, we are now in position to discuss the generalization of the notion of pseudo-differential operator to compact Riemannian manifolds without boundary. Let $(\mathcal{M}, g)$ be such a manifold and let $E$ and $F$ be smooth complex vector bundles on $\mathcal{M}$ equipped with Hermitian inner products $\langle \cdot, \cdot \rangle_E$ and $\langle \cdot, \cdot \rangle_F$.

In the manifold case, unlike on $\mathbb{R}^N$, it is necessary to introduce infinitely smoothing operators before defining pseudo-differential ones, as the $\sim$ equivalence relation is used to transfer operators from charts in $\mathcal{M}$ to coordinate patches on $\mathbb{R}^N$. Those operators are defined in a way completely analogous to the one on $\mathbb{R}^N$. $P : C^\infty(E) \to C^\infty(F)$ is said to be infinitely smoothing if, for any real $s$ and $m$, it extends to a bounded linear map $P : H_s(E) \to H_{s+m}(F)$. The space of such operators is denoted $\Psi^{-\infty}(E, F)$.

**Definition 2.2.9** ($\Psi$DOs on Manifolds). *Let $P : C^\infty(E) \to C^\infty(F)$ be a linear map. $P$ is called a pseudo-differential operator of order $m$ if, modulo infinitely smoothing operators, it can expressed as a finite sum*

$$P = \sum_i P_i \quad mod(\Psi^{-\infty}(E, F)) \tag{2.59}$$

*where for each $P_i$ there exists a bundle trivializing chart $U_i$ equipped with a system of local coordinates $x_i : U_i \to \mathbb{R}^N$ in which it can be expressed as a pseudo-differential operator of order $m$ with compact support in $x_i(U)$. The space of such operators is denoted $\Psi^m(E, F)$. The space of operators of all orders is denoted $\Psi(E, F) = \cup_m \Psi^m(E, F)$. For $P, Q \in \Psi(E, F)$ it is said that $P \sim Q$ if $P - Q \in \Psi^{-\infty}(E, F)$.*

**Remark 2.2.7.** *The sum in the previous definition allows one to build an operator on a manifold out of operators defined on coordinate patches of (possibly disjoint) charts. This is quite different from the way one normally defines differential operators on manifolds. In order to bridge this gap, consider a differential operator $P$ defined through the usual means. In order to make it conform to the above definition, choose a finite partition of unity $\{u_i\}_{i=1}^r$. The finite sum form of $P$ then becomes*

$$P = \sum_{i=1}^r u_i P \tag{2.60}$$

26

*In essence, the construction in Definition 2.2.9 reverses those steps.*

Results obtained for operators on smooth bundles of $\mathbb{R}^N$ generalize to the manifold setting via the construction of Definition 2.2.9.

**Theorem 2.2.1.** *Let $E, F$ and $G$ be smooth vector bundles over a compact Riemannian manifold $(\mathcal{M}, g)$. Let $P \in \Psi^m(E, F)$ and $Q \in \Psi^l(F, G)$. The following hold:*

*(a) $P$ extends to a bounded linear map $P : H_s(E) \to H_{s-m}(F)$ for all $s \in \mathbb{R}$*

*(b) $Q \circ P \in \Psi^{l+m}(E, G)$*

*(c) $P^* \in \Psi^m(F, E)$*

*(d) A diffeomorphism $\phi : \mathcal{M} \to \mathcal{M}$ induces a push-forward of operators via $(\phi_* P) = \phi_* \circ P \circ \phi^*$*

We are not going to generalize the notion of symbols for pseudodifferential operators on compact Riemannian manifolds. The main reason for this omission is that we are only really interested in differential operators, for which we already have a notion of principal symbol. In order to globally define a notion of symbol of a pseudodifferential operator, one can introduce an interesting generalization of a Fourier transform, valid for functions of manifolds. See [54] for further details.

## 2.3 Elliptic Operators

This section is where the work done above starts to pay off, as we finally introduce elliptic operators and obtain a spectral theorem for them. The proofs of this section are more detailed as they showcase the usefulness of prior definitions and results. Once again we closely follow [54]. We start by defining elliptic differential operators on compact Riemannian manifolds, and then generalize the definition to pseudo-differential ones.

**Definition 2.3.1** (Elliptic PDO)**.** *Let $P : E \to F$ be a differential operator of order $m$ over a Riemannian manifold $(\mathcal{M}, g)$. $P$ is said to be elliptic if for each non-zero tangent vector $\xi \in T^*\mathcal{M}$, its principal symbol $\sigma_\xi(P) : E_x \to F_x$ is invertible.*

For pseudo-differential operators, the corresponding definition needs to be built from the properties of the principal symbol on coordinate patches in $\mathbb{R}^N$ that correspond to compact charts. Consequently, ellipticity in pseudo-differential operators is first defined for operators on $\mathbb{R}^N$.

**Definition 2.3.2** (Elliptic $\Psi$DO; $\mathbb{R}^N$)**.** *Let $P \in \Psi^m$ have symbol $p$. $P$ is said to be elliptic if there exists a constant $C > 0$ such that for all $|\xi| \geq C$, the inverse matrix of $p(x, \xi)$ exists and obeys the following inequality:*

$$\left| p(x, \xi)^{-1} \right| \leq C(1 + |\xi|)^{-m} \tag{2.61}$$

**Remark 2.3.1.** *Differential operators elliptic in the sense of Definition 2.3.1 are also elliptic in the sense of Definition 2.3.2. Recall that the full symbol of a differential operator is a polynomial expression in $\xi$ whose highest order terms constitute the principal symbol. Thus, for large enough $|\xi|$, the behaviour of the symbol is dominated by the principal symbol, which is invertible for all $|\xi| > 0$. Thus, the full symbol is invertible for large enough $|\xi|$. The inequality is established by methods similar to those used to prove equivalence of different norms on Sobolev spaces.*

**Remark 2.3.2.** *There exist elliptic operators of all orders. Let $m \in \mathbb{R}$ and set $p = (1 + |\xi|^2)^m \mathbb{1}$. Manifestly, $p$ is the symbol of an elliptic operator $P \in \Psi^m$ [25].*

The power of elliptic operators resides in their invertibility up to infinitely smoothing operators.

**Theorem 2.3.1** (Parametrix; $\mathbb{R}^N$)**.** *Let $P \in \Psi^m$ be elliptic. There exists $Q \in \Psi^{-m}$ such that*

$$\begin{aligned} PQ &= \mathbb{1} - S' \\ QP &= \mathbb{1} - S \end{aligned} \tag{2.62}$$

*where $S, S'$ are infinitely smoothing. $Q$ is unique up to equivalence and is said to be a parametrix for $P$.*

*Proof.* Our proof follows [54]. The overall strategy is to build an asymptotic development for the operator $Q$, proceeding inductively order by order in the formal sum. Denote this construction in the following way

$$q \sim \sum_{k}^{\infty} q_k \tag{2.63}$$

where $q_0$ is taken to be of order $-m$, $q_1$ of order $-m - 1$ etc. Then, Lemma 2.2.5 can be used to construct $Q$. Recall that the order of $Q$ will be that of the highest order symbol in the formal expansion, $-m$ in this case.

Let $p$ be the symbol of $P$ and let $C$ denote the constant of equation (2.61). As in the proof of Lemma 2.2.5, we need to introduce a selection function. Let $\chi : \mathbb{R}^+ \to [0,1]$ be a smooth function such that $\chi(t) = 0$ for $t \leq C$ and $\chi(t) = 1$ for $t \geq 2C$. Define

$$q_0 = \begin{cases} 0, & \text{if } |\xi| \leq C \\ \chi(|\xi|)p(x,\xi)^{-1}, & \text{otherwise} \end{cases} \tag{2.64}$$

First, it is necessary to prove that $q_0$ is indeed of order $-m$. By Definition 2.2.4, there must exist constants $C_{\alpha,\beta}$ such that:

$$\left| D_x^\alpha D_\xi^\beta q_0 \right| \leq C_{\alpha,\beta}(1 + |\xi|)^{-m-|\beta|} \quad \forall \alpha, \beta \tag{2.65}$$

This bound is trivially established for $|\xi| \leq C$. We thus concentrate our attention on $\xi > C$. For $\alpha = \beta = 0$, the existence of $C_{\alpha=0,\beta=0}$ is established by the bound on the inverse symbol of Definition 2.3.2. For other $\alpha, \beta$ we proceed inductively. First, set $\alpha = 0$. That is, ignore the $x$ derivatives. Using $pp^{-1} = \mathbb{1} = p^{-1}p$ and the Leibniz rule it is straightforward to show that

$$\partial_{\xi_i} p^{-1} = -p^{-1}(\partial_\xi p)p^{-1} \tag{2.66}$$

By the bound on $p^{-1}$ of Definition 2.3.2 and the bounds on derivatives of $p$ of Definition 2.2.4, there exists a constant $K_{0,1}$ such that the following bound holds:

$$\left| \partial_{\xi_i} p^{-1} \right| = \left| p^{-1}(\partial_{\xi_i} p)p^{-1} \right| \leq C^2 K_{0,1}(1 + |\xi|)^{-2m}(1 + |\xi|)^{m-1} \\ = C_{0,1}(1 + |\xi|)^{-m-1} \tag{2.67}$$

Bounds on higher order $\xi$ derivatives are obtained in a similar fashion. By the Leibniz rule, successive derivatives of Equation 2.66 take the form of sums of products of different order derivatives of $p$ and $p^{-1}$. For $\beta$ successive derivatives, the total order of differentiation of any given term of the sum is $|\beta|$. For each term in the sum, a bound proportional to $(1 + |\xi|)^w$ is obtained. The whole sum is then bounded by the term corresponding to the highest $w$, multiplied by some constant. For all $|\beta|$ this is achieved when all derivatives are on $p$. That is, the term with the highest $w$ is of the form $p^{-1}(d_\xi^\beta p)p^{-1}$. As above, one has the following bound:

$$\left| p^{-1}(d_\xi^\beta p)p^{-1} \right| \leq C_{0,|\beta|}(1 + |\xi|)^{-m-|\beta|} \tag{2.68}$$

29

where $C_{0,|\beta|}$ is chosen large enough to bound all of $d_\xi^\beta p^{-1}$. Up to a change in the constant $C_{0,\beta}$, this bound remains true for $q_0$, as all derivatives of $\chi(|\xi|)$ are clearly bounded.

The bounds established for $\alpha = 0$ already have the desired form. It remains thus to show that allowing $\alpha$ to be nonzero does not disrupt them. It is so for bounds on the derivatives of $p$, by Definition 2.2.4. Notice that Equation (2.66) holds if $\xi$ derivatives are replaced with $x$ derivatives. By the same argument as above, the following bound is obtained:

$$\left| \partial_{x_i} p^{-1} \right| \leq C_{0,1}(1 + |\xi|)^{-m} \tag{2.69}$$

Since $x$ derivatives do not modify the bounds on $d_\xi^\beta p$ nor on $\partial_{x_i} p^{-1}$, one then inductively shows that bounds on $|D_x^\alpha D_\xi^\beta q_0|$ are the same as those on $|D_\xi^\beta q_0|$. Thus, $q_0$ is indeed a symbol of order $-m$.

Let $Q_0$ be the operator associated to the symbol $q_0$. By the rule for symbol composition (Lemma 2.2.8), we have

$$
\begin{aligned}
sym(Q_0 P - \mathbb{1}) &= \sum_\alpha \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_0)(D_x^\alpha p) - 1 \\
&= (\chi(|\xi|) - 1) + \mathcal{R}_0
\end{aligned}
\tag{2.70}
$$

Since there exists a constant $C_{\chi,w}$ such that $|\chi(|\xi|) - 1| \leq C_{\chi,w}(1 + |\xi|)^w$ for all $w$, $(\chi(|\xi|) - 1)$ is infinitely smoothing. The remainder $\mathcal{R}_0$ is thus the symbol of $(Q_0 P - \mathbb{1})$, up to equivalence. The order of $\mathcal{R}_0$ is that of the $|\alpha| = 1$ terms, which is readily shown to be $-1$.

This process now needs to be repeated for all orders. Let $Q_1$ be such that

$$sym((Q_0 + Q_1)P - \mathbb{1}) = (\chi(|\xi|) - 1) + sym(Q_1 P) + \mathcal{R}_0 \in S^{-2} \tag{2.71}$$

In other words, $Q_1$ must be such that the symbol of $Q_1 P$ cancels the highest order part in $\mathcal{R}_0$, up to equivalence. Let $q_1$ be the symbol of $Q_1$. Then, by the law of composition of symbols,

$$
\begin{aligned}
sym(Q_1 P) + \mathcal{R}_0 &= \sum_{|\alpha|} \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_1)(D_x^\alpha p) + \sum_{|\alpha| \geq 1} \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_0)(D_x^\alpha p) \\
&= q_1 p + \sum_{|\alpha| \geq 1} \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_1)(D_x^\alpha p) + \sum_{|\alpha| \geq 1} \frac{i^{|\alpha|}}{\alpha!} (D_\xi^\alpha q_0)(D_x^\alpha p)
\end{aligned}
\tag{2.72}
$$

30

One must now cleverly pick a $q_1$ that would cancel the highest order terms in the above expression, up to equivalence. Notice that if the order of $q_1$ is low enough, the first summation will not contain leading order terms. Thus, it is only necessary to cancel the leading order terms in the second sum. This can be done by setting:

$$q_1 = -\left[\sum_{|\alpha|=1} \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha q_0)(D_x^\alpha p)\right] q_0 \in S^{-m-1} \tag{2.73}$$

This indeed yields, up to equivalence,

$$sym(Q_1 P) + \mathcal{R}_0 = \sum_{|\alpha|\geq 1} \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha q_1)(D_x^\alpha p) + \sum_{|\alpha|\geq 2} \frac{i^{|\alpha|}}{\alpha!}(D_\xi^\alpha q_0)(D_x^\alpha p) \in S^{-2} \tag{2.74}$$

Iterating this discussion leads to the following formula for $q_k$

$$q_k = -\sum_{j=0}^{k-1}\left[\sum_{|\alpha|+j=k} (D_\xi^\alpha q_j)(D_x^\alpha p)\right] q_0 \in S^{-m-k} \tag{2.75}$$

This completes the construction of the formal development of $q$. Lemma 2.2.5 ensures that it leads to an operator $Q \in \Psi^{-m}$, unique modulo infinitely smoothing operators. To show that $QP - \mathbb{1}$ is indeed infinitely smoothing, it is permissible to select a particular representative of $Q$, as the composition of any pseudo-differential operator with an infinitely smoothing operator is also infinitely smoothing. For instance, one can use Lemma 2.2.5 to explicitly construct the symbol of $Q$. By the preceding discussion and the rule of composition of symbols, it then becomes clear that $sym(PQ - \mathbb{1}) \in S^\infty$ and thus that $PQ - \mathbb{1} \in \Psi^\infty$.

Similarly, one can build $Q'$ such that $PQ' - \mathbb{1}$ is infinitely smoothing. It is permissible (and usually desirable) to identify $Q$ and $Q'$, since.

$$Q \sim Q(PQ') = (QP)Q' \sim Q' \tag{2.76}$$

$\square$

The above result now needs to be generalized to compact Riemannian manifolds without boundary. Notice that we never defined the notion of ellipticity of a pseudo-differential

operator on a compact manifold, nor did we define the total symbol of a pseudo-differential operator on a manifold. Consequently, we limit our discussion to differential operators, for whom those definitions have been stated. This loss of generality is of little consequence to our endeavours, as we are mostly interested in Laplacians. Nonetheless, the results still hold for pseudo-differential operators. A reader interested in the full picture is encouraged to consult [25].

**Theorem 2.3.2.** *Let $E, F$ be smooth vector bundles over a compact, boundaryless Riemannian manifold $(\mathcal{M}, g)$. Let $P : C^\infty(E) \to C^\infty(F)$ be a differential operator of order $m$. Then there exists an operator $Q \in \Psi^{-m}(F, E)$ such that*

$$
\begin{aligned}
PQ &= \mathbb{1} - S' \\
QP &= \mathbb{1} - S
\end{aligned}
\tag{2.77}
$$

*where $S$ and $S'$ are infinitely smoothing operators.*

*Sketch of Proof.* The proof relies on the choice of a good presentation of the bundles $E$ and $F$ with coordinate charts $\{U_j\}_j$ and of a suitable partition of unity $\{\psi_i\}_i$. The operator defined by the action of $P$ restricted to the chart corresponding to $U_j$ is denoted $P_j$. Let $Q_j$ be its parametrix and set $S_j = \mathbb{1} - Q_j P_j$. The global parametrix $Q$ and the global remainder $S$ are thus

$$
\begin{aligned}
Q &= \sum_j \psi_j Q_j \\
S &= \sum_j \psi_j S_j
\end{aligned}
\tag{2.78}
$$

See [54] for further details. □

The parametrix allows one to show some key properties of elliptic differential operators.

**Theorem 2.3.3** (Smoothness)**.** *Let $P : C^\infty(E) \to C^\infty(F)$ be an elliptic differential operator over a compact Riemannian manifold $(\mathcal{M}, g)$. For any open $U \subset \mathcal{M}$, and any $u \in H_s(E)$,*

$$
Pu|_U \in C^\infty \implies u|_U \in C^\infty
\tag{2.79}
$$

*In particular, if $E = F$ and $Pu = \lambda u$ for some $\lambda \in \mathbb{C}$, then $u$ is smooth.*

*Proof.* Let $Q$ be a parametrix of $P$ such that $\mathbb{1} = QP + S$, with infinitely smoothing $S$. Clearly, $u = \mathbb{1}u = QPu + Su$ is smooth on $U$ as $Pu$ and $Su$ are both smooth and pseudo-differential operators map smooth functions into smooth functions [54].

Since $P$ is a differential operator, it is of positive order. $\lambda\mathbb{1}$ is thus of lower order and $P - \lambda\mathbb{1}$ retains its ellipticity. By hypothesis, $(P - \lambda\mathbb{1})u = 0$, which establishes the smoothness of $u$. □

**Remark 2.3.3.** *The above theorem establishes the crucial fact that the eigenfunctions of elliptic differential operators are smooth. This is also true for elliptic pseudo-differential operators of positive order.*

Another powerful result that can be proven via the existence of the parametrix is the Fundamental Elliptic Estimate.

**Theorem 2.3.4** (Fundamental Elliptic Estimate)**.** *Let $P : C^\infty(E) \to C^\infty(F)$ be an elliptic differential operator of order $m$ over a compact Riemannian manifold $(\mathcal{M}, g)$. For each $s$ there exists a constant $C_s$ such that*

$$\|u_s\| \leq C_s(\|u\|_{s-m} + \|Pu\|_{s-m}) \ \forall \ u \in H_s(E) \tag{2.80}$$

*Proof.* Let $Q$ be a parametrix of $P$ such that $\mathbb{1} = QP + S$, with infinitely smoothing $S$. One can thus estimate:

$$\|u\|_s = \|QPu + Su\|_s \leq \|QPu\|_s + \|Su\|_s \leq C(\|Pu\|_{s-m} + \|u\|_{s-m}) \tag{2.81}$$

The last inequality is obtained from the fact that $Q$ is of order $-m$ and $S$ can also be considered of order $-m$. Lemma 2.2.3 provides the required inequality. □

**Remark 2.3.4.** *Jointly, Lemma 2.2.3 and the Fundamental Elliptic Estimate show that the norms $\|\cdot\|_s$ and $\|P\cdot\|_{s-m} + \|\cdot\|_{s-m}$ are equivalent.*

Recall that by Lemma 2.2.4, infinitely smoothing operators are compact. Operators on Hilbert spaces invertible modulo compact operators are said to be Fredholm. An alternative characterization of Fredholm operators is that the kernel and cokernel (quotient of the target Hilbert space by the closure of the range) of a Fredholm operator are finite-dimensional and its range is closed. The existence of the parametrix ensures that any elliptic $P : C^\infty(E) \to C^\infty(F)$ of order $m$ extends to a Fredholm map $P : H_s(E) \to H_{s-m}(F)$ [54]. For formally self-adjoint operators, this has the following consequence:

**Theorem 2.3.5.** *Let* $P : C^\infty(E) \to C^\infty(E)$ *be an elliptic differential operator over a compact Riemannian manifold without boundary that is formally self-adjoint with respect to the $L_2$ inner product. Then, the following $L_2$-orthogonal decomposition holds:*

$$C^\infty(E) = ker P \oplus im P \qquad (2.82)$$

*Proof.* See [54]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

**Theorem 2.3.6** (Spectral Resolution). *Let* $P : C^\infty(E) \to C^\infty(E)$ *be an elliptic differential operator of order $m$ over a compact Riemannian manifold without boundary that is formally self-adjoint with respect to the $L_2$ inner product. Then $P$ has a discrete spectral resolution $\{\psi_n, \lambda_n\}_{n=1}^{\infty}$ with $\psi_n \in C^\infty(E)$ and $\lambda_n$ real. The eigenspaces of $P$ have finite dimension. The spectrum of $P$ has no finite accumulation points.*

*Proof.* The proof follows [25]. By the above discussion, $P : H_m(E) \to L_2(E)$ is Fredholm. Let $\tilde{P}$ be a restriction of $P$ such that:

$$\tilde{P} := P : (ker P)^\perp \cap H_m(E) \longrightarrow (ker P)^\perp \cap L_2(E) \qquad (2.83)$$

$\tilde{P}$ is an isomorphism, since the above restriction excises both the kernel and cokernel of $P$. Thus, one can construct an inverse $\tilde{Q}$ of $\tilde{P}$. $\tilde{Q}$ can then be extended to act upon all of $L_2(E)$ by setting its action to be zero on the finite-dimensional space $(ker P)$. $\tilde{Q}$ is manifestly a choice of parametrix and thus is a pseudo-differential operator of order $-m$. Consequently, the map $\tilde{Q} : L_2(E) \to H_m(E)$ is bounded. This map may then be composed with the natural projection of $H_m(E)$ onto $L_2(E)$, which is compact as $m > 0$. The overall map $Q : L_2(E) \to L_2(E)$ is thus compact, as the composition of a bounded linear map and a compact map is compact [57].

It is straightforward to show from the decomposition $C^\infty(E) = ker P \oplus im P$ that the formal self-adjointness of $P$ ensures that of $Q$. $Q$ can then be extended to be self-adjoint on $L_2(E)$. The spectral theorem for compact self-adjoint operators then applies, providing $Q$ with a discrete spectral resolution $\{\psi_n, \mu_n\}_{n=1}^{\infty}$, with $\mu_n$ real, of finite multiplicity and having no non-zero accumulation points [57]. The spectral resolution for $P$ is obtained by inverting the nonzero eigenvalues:

$$\lambda_n = \begin{cases} \mu_n^{-1}, & \text{if } \mu_n \neq 0 \\ 0, & \text{if } \mu_n = 0 \end{cases} \qquad (2.84)$$

Compare the above expression with the definition of the Moore-Penrose pseudoinverse given in Appendix A. Since $S$ and $P$ share their kernel, $S\psi_n = 0$ implies $P\psi_n = 0$. Similarly,

when $\mu_n \neq 0$, $S\psi_n = \mu_n\psi_n$ implies $P\psi_n = PS\mu_n^{-1}\psi = \lambda_n\psi_n$. Thus, $\{\psi_n, \lambda_n\}$ is indeed a spectral resolution for $P$ with eigenspaces of finite dimension and eigenvalues with no finite accumulation point. Smoothness of the $\psi_n$ is established by Theorem 2.3.3. $\qquad \square$

**Remark 2.3.5.** *The above theorem also holds for compact Riemannian manifolds with sufficiently regular boundary, provided that adequate boundary conditions are imposed. The theory of boundary conditions for pseudo-differential operators is a subtle subject that is beyond the scope of this work. For manifolds with smooth boundaries, the interested reader is referred to [32, 66, 67]. For our purposes, it is sufficient to say that Dirichlet or Neumann boundary conditions are indeed appropriate. Much more exotic boundary conditions are discussed in [26].*

The last important piece of information it remains to obtain about the spectrum of $P$ is the growth properties of its eigenvalues. Since $P$ is a differential operator, it is unbounded, so its eigenvalues tend to infinity in some sense. The following theorem gives a bound on the growth of the eigenvalues:

**Theorem 2.3.7.** *Let $P$ and $\{\psi_n, \lambda_n\}_{n=1}^{\infty}$ be as in Theorem 2.3.6. Let $E_\lambda$ denote the eigenspace corresponding to an eigenvalue $\lambda$. Define the following dimension counting function:*

$$d(\Lambda) = dim \left( \bigoplus_{|\lambda| \leq \Lambda} E_\lambda \right) \tag{2.85}$$

*Then, there exists a constant $C$ such that*

$$d(\Lambda) \leq C\Lambda^{N(N+2m+2)/2m} \tag{2.86}$$

*Proof.* See [54]. It is an application of the fundamental elliptic estimate. $\qquad \square$

$P$ is said to be non-negative if $(Pu, u)_{L_2} \geq 0$ for all smooth $u$. Consequently, the eigenvalues of $P$ are non-negative. The estimates on the growth of the eigenvalues then take a much simpler form:

**Theorem 2.3.8.** *Let $P$ and $\{\psi_k, \lambda_k\}_{k=1}^{\infty}$ be as in Theorem 2.3.6. Suppose that $P$ is non-negative. Then, its spectrum $\{\lambda_k\}_{k=1}^{\infty}$ can be ordered as follows:*

$$0 \leq \lambda_1 \leq \lambda_2 \leq ... \rightarrow \infty \tag{2.87}$$

35

*In this eigenvalue numbering scheme, there exists a constant $C$ such that, for all $k$*

$$\lambda_k \geq C k^{2m/N(N+2m+2)} \tag{2.88}$$

*Proof.* For a non-degenerate spectrum, $d(\lambda_k) = k$ and the result follows from Theorem 2.3.7. For a degenerate spectrum, the same reasoning holds for the highest $k$ in an eigenspace. The remainder of the eigenspace has lower bounds, so the result still holds. $\quad\square$

**Remark 2.3.6.** *The last bound is used in the next section to establish the convergence of the heat kernel, an important tool in spectral geometry.*

## 2.4   Heat Equation and Heat Kernel

The Heat Equation is one of the classical equations of mathematical physics. On $\mathbb{R}^N$ it takes the form

$$\frac{\partial}{\partial t} u + \Delta u = 0 \tag{2.89}$$

where $\Delta$ is the positive definite Laplacian on $\mathbb{R}^N$ and $t$ is a time coordinate taken to vary over $(0, \infty)$. Physically, this equation describes the conduction of heat in a slab of material. Thus, the problem is normally defined over a bounded domain of $\mathbb{R}^N$ and boundary conditions motivated by the physical situation are imposed. Moreover, an initial condition $\lim_{t \to 0^+} u = f$, representing an initial distribution of temperature, is supplied. The equation is then usually solved by separation of variables. First, one postulates exponential behaviour in the time variable, eliminating time dependence. Then, one solves the spatial part as an eigenvalue problem for the Laplacian. The solution for the initial value problem is a linear combination of the eigenfunctions of the Laplacian with coefficients that vary exponentially in time. This classical approach is exposed in most introductory textbooks on partial differential equations, say [30].

The above discussion generalizes to the case of non-negative elliptic differential operators on manifolds.

**Definition 2.4.1** (Heat Equation). *Let $P : C^\infty(E) \to C^\infty(E)$ be a non-negative formally self-adjoint elliptic differential operator on a compact Riemannian manifold. If the manifold has boundaries, appropriate boundary conditions are assumed to hold so that Theorem*

*2.3.6 applies. The following system of equations is called the heat equation for $P$:*

$$(\partial_t + P)u(x,t) = 0$$
$$\lim_{t \to 0^+} u(x,t) = f(x) \qquad (2.90)$$

**Definition 2.4.2** (Heat Operator). *Let $P$ be as in Definition 2.4.1. The heat operator of $P$ is defined to be $e^{-tP}$. In terms of a spectral resolution $\{\psi_k, \lambda_k\}_{k=1}^{\infty}$ of $P$, this operator takes the form*

$$e^{-tP}f(x) = \int_{\mathcal{M}} K_t(x,y)f(y)dV_y \qquad (2.91)$$

*where the heat kernel $K_t(x,y)$ is given by*

$$K_t(x,y) = \sum_{k=1}^{\infty} e^{-\lambda_k t} \psi_k(x) \otimes \psi_k^*(y) \qquad (2.92)$$

**Lemma 2.4.1.** *For any closed interval $I \in (0, \infty)$ and any $r \geq 0$, the series in equation (2.92) converges uniformly in the $C^r$ topology on $I \times \mathcal{M} \times \mathcal{M}$. Thus $K_t(x,y)$ is smooth.*

*Sketch of Proof.* Sobolev's theorem 2.1.1 and elliptic estimates are used to bound the $C^r$ norm of the partial sums. This bound depends upon $\{\lambda_k\}_k$. Theorem 2.3.8 ensures that the spectrum grows quickly enough so that the exponential factors enforce convergence. See [54] for details. $\qquad \square$

**Theorem 2.4.1.** *For a fixed $t$, the heat operator is infinitely smoothing. Also, $u(x,t) = e^{-tP}f(x)$ satisfies the heat equation and is smooth for any $f \in L_2(E)$.*

*Proof.* The heat operator is infinitely smoothing as it is an integral operator with smooth kernel. The smoothness of $u(x,t)$ is then guaranteed. Taking a time derivative of $K_t(x,y)$, one gets

$$\partial_t K_t(x,y) = \sum_{k=1}^{\infty} -\lambda_k e^{-\lambda_k t} \psi_k(x) \otimes \psi_k^*(y)$$
$$= -P_x K_t(x,y) \qquad (2.93)$$

Once multiplied by $f$ and integrated, this yields $\partial_t u = -Pu$. $\qquad \square$

**Remark 2.4.1.** *The heat operator $e^{-tP}$ was explicitly constructed to solve the heat equation. This was made possible by the spectral resolution of $P$, whose existence and smoothness took a great deal of work to prove. Interestingly, one can proceed in the opposite direction, starting with the definition of the heat equation for $P$, constructing a heat operator and then proving that $P$ has a spectral resolution of smooth sections. For the Laplace-Beltrami operator, this approach is discussed in [13].*

A particularly interesting object is the $L_2$-trace of the heat kernel. In later chapters it will be used as the main tool to extract geometric information from the spectrum of various Laplacians.

**Definition 2.4.3.** *Let $P$ be as in Definition 2.4.1. The trace of the heat kernel for $P$ is defined to be*

$$
\begin{aligned}
Tr(e^{-tP}) &= \int_{\mathcal{M}} trace_x[K_t(x,x)]dV_x \\
&= \sum_{k=1}^{\infty} e^{-\lambda_k t}
\end{aligned}
\tag{2.94}
$$

*The heat kernel is analytic for all $t > 0$ [54].*

Notice that the heat kernel of $P$ is uniquely determined by the spectrum of $P$. It is thus a spectral invariant. Its asymptotic expansion as $t \downarrow 0$, which we are about to define, contains a great deal of geometric information. In our discussion of the asymptotics of the heat kernel, we follow [25].

**Definition 2.4.4** (Asymptotic Expansion). *Let $t > 0$ and let $f_i(x), h(x,t) \in End(E)$. $h(x,t)$ is said to have an asymptotic expansion*

$$
h(x,t) \sim \sum_i f_i(x) t^{(i-N)/m}
\tag{2.95}
$$

*if for every $k$, there exist $M = M(k)$ and constants $C_k$ such that*

$$
\left\| h(x,t) - \sum_{i \leq M} f_i(x) t^{(i-N)/m} \right\|_{\infty,k} \leq C_k t^k
\tag{2.96}
$$

*If it exists, the asymptotic expansion is unique. Note that the series needs not to converge.*

**Remark 2.4.2.** *The above definition is specialized for the task at hand. For general asymptotic expansions, one would absorb the common factor of $t^{-N/m}$ into the definition of the $f_i$.*

**Theorem 2.4.2** (Heat Kernel Coefficients). *The Heat Kernel of a formally self-adjoint elliptic differential operator of order d has the following asymptotic expansion:*

$$Tr(e^{-tP}) \sim \sum_n a_n(P) t^{(n-N)/d} \tag{2.97}$$

*The coefficients $a_n$ depend solely on the operator $P$.*

*Sketch of Proof.* The full discussion leading to the proof of the theorem is quite lengthy. We refer the interested reader to [25]. The main complication is that one must obtain a generalized version of the symbol calculus valid for operators depending upon a complex parameter $\lambda$ (see [25] or [69]). The goal is to study the resolvent $(P - \lambda)^{-1}$ of $P$. The resolvent is constructed from successive approximations, similarly to how the parametrix was obtained in the proof of Theorem 2.3.1. The successive approximations of the resolvent are then used to obtain asymptotic expansions of the heat kernel evaluated on the diagonal $x = y$, at which point the asymptotic coefficients still have $x$ dependence. Taking the $L_2$ trace of the kernel on the diagonal removes all dependence upon $x$ and yields the asymptotic expansion for $Tr(e^{-tP})$. □

**Remark 2.4.3.** *Another path to this asymptotic expansion is taken in [13] for the Laplace-Beltrami operator. There, as mentioned earlier, the spectral resolution of $\Delta$ is shown to exist from the existence of the heat operator. In order to show the existence of the heat operator for Riemannian manifolds, [13] first constructs the heat operator on $\mathbb{R}^N$ and then obtains the one on $(\mathcal{M}, g)$ by successive approximations. As a by-product of this construction, the asymptotic expansion of the trace of the heat kernel is obtained.*

We shall not touch upon techniques involved in the computation of the heat trace coefficients $a_n$ beyond the few following facts. On a compact manifold without boundary, all odd $a_n$ vanish. On a manifold with boundary, the $a_n$ will be generically composed of contributions from both the operator $P$ and the boundary conditions. For odd $n$, only the boundary terms contribute. The $a_n$ are computable from the symbol of $P$ and the boundary conditions, when appropriate [25, 31].

Asymptotic expansion coefficients can be explicitly computed for a particular class of operators called Laplace-type operators.

**Definition 2.4.5** (Laplace Type Operators). *A differential operator $P$ acting on sections of some smooth vector bundle $E$ of a compact Riemannian manifold $(\mathcal{M}, g)$ is said to be of Laplace type if its leading symbol is given by*

$$\sigma_\xi(P) = \|\xi\|^2 = g^{ij} \mathbb{1}_E \xi_i \xi_j \tag{2.98}$$

39

*where $g^{ij}$ denotes the inverse metric expressed in a local coordinate basis and the Einstein summation convention is used.*

Note that since the principal symbol of a Laplace type operator is given by the metric, the asymptotic expansion of its heat trace is expected to contain at least some geometric information about the Riemannian manifold $(\mathcal{M}, g)$. The remainder of this thesis is dedicated to the extraction of geometric information from operators of Laplace type.

# Chapter 3

# Laplace Operators

Inverse spectral geometry is built around the relationship between spectra of various Laplacians and the geometry of Riemannian manifolds, with the odd Dirac operator sometimes thrown into the mix. We shall limit our discussion to Laplace operators. Most of that information comes from the coefficients in the asymptotic expansion of the heat trace. In this fashion, different Laplacians yield different geometric information about the manifold. First, we define the most commonly used Laplacians and showcase basic symbol computation techniques by computing their leading symbols. As promised by the terminology introduced at the very end of the previous chapter, Laplacians will all turn out to be of Laplace type. Then, we review the geometric information one can extract from their spectra via known methods.

## 3.1 Laplacians

In this section, we review the definitions of various Laplacians on Riemannian manifolds. In truth, only the Laplace-Beltrami operator and Hodge Laplacian are of immediate interest to our endeavours, as they are the ones most often used in spectral geometry. The definition of the covariant Laplacian is included for the sake of completeness. We will use the symbol $\Delta$, possibly with subscripts, exclusively to denote the various Laplacians of this chapter.

For the remainder of this chapter, let $(\mathcal{M}, g)$ be an oriented $N$-dimensional Riemannian manifold. We take $|g|$ to be the absolute value of the determinant of the metric. The volume form associated to the metric $g$ is denoted $dV_g$. The Levi-Civita connection compatible with $g$ will be denoted by $\nabla$. The Lie derivative with respect to a vector field $X \in C^\infty(T\mathcal{M})$

will be written $\mathcal{L}_X$. For details on all matters differential geometric, we refer the reader to [72] and [59], in increasing order of complexity. The discussion below is mainly inspired by [59].

### 3.1.1 Laplace-Beltrami Operator

The Laplace-Beltrami operator is a straightforward generalization of the usual Laplacian, familiar to the reader from vector calculus [30]. Recall that the Laplacian from vector calculus is defined as the divergence of the gradient of a function:

$$\Delta f = \mathbf{div}(\mathbf{grad} f) = \partial_1^2 + ... + \partial_N^2 \tag{3.1}$$

This operator comes out to be non-positive definite, rather than non-negative, which is a slight technical inconvenience as we prefer working with non-negative spectra. Thus, when constructing the Laplace-Beltrami operator, one really generalizes the negative of the usual Laplacian. This is sometimes referred to as the geometer's Laplacian.

The construction of the Laplace-Beltrami operator requires only a generalization of the notions of gradient and divergence to manifolds. Recall that the gradient of $f \in C^\infty(\mathcal{M})$ is a tangent vector field $\mathbf{grad} f \in C^\infty(T\mathcal{M})$ that encodes the directional derivative of $f$ in the direction of any $X \in C^\infty(TM)$ through contraction with the metric. This is expressed as follows:

$$g(\mathbf{grad} f, X) = Xf \tag{3.2}$$

In turn, the divergence of a vector field $X \in C^\infty(T\mathcal{M})$ is defined to be the unique function $\mathbf{div} X \in C^\infty(M)$ such that

$$(\mathbf{div} X)dV_g = \mathcal{L}_X dV_g \tag{3.3}$$

The Laplace-Beltrami operator is then defined to be the negative of the divergence of the gradient, mapping $C^\infty(\mathcal{M})$ into itself. In local coordinates $\{x_i\}_{i=1}^N$, the action of $\Delta$ upon a function $f \in C^\infty(\mathcal{M})$ takes the form

$$\Delta f = -\frac{1}{\sqrt{|g|}}\partial_i\left(\sqrt{|g|}g^{ij}\partial_j f\right) \tag{3.4}$$

For the Euclidean metric on $\mathbb{R}^N$, it indeed reduces to the negative of the classical sum of second derivatives Laplacian. Moreover, note that it is indeed an operator of Laplace type, as the leading order terms are

$$\Delta = -g^{ij}\partial_i\partial_j + ... \tag{3.5}$$

The leading symbol can be read off directly $\sigma_\xi(\Delta) = g^{ij}\xi_i\xi_j = |\xi|^2$.

### 3.1.2 Hodge Laplacian

The Hodge Laplacian is a generalization of the Laplace-Beltrami operator that acts upon differential $p$-forms. The space of smooth $p$-forms (i.e. smooth sections of $\Lambda^p(T^*\mathcal{M})$) will be denoted $\Omega^p(\mathcal{M})$. The exterior algebra of smooth forms of all orders is denoted

$$\Omega^\bullet(\mathcal{M}) = \bigoplus_{p\geq 0} \Omega^p(\mathcal{M}) \tag{3.6}$$

Let $d$ denote the exterior derivative $d : \Omega^p(\mathcal{M}) \to \Omega^{p+1}(\mathcal{M})$ and let $* : \Omega^p \to \Omega^{N-p}$ denote the Hodge star operator. It is well known that the dual of $d$ with respect to the $L_2$ inner product, called the codifferential, can be defined as

$$\begin{aligned} \delta &: \Omega^p(\mathcal{M}) \to \Omega^{p-1}(\mathcal{M}) \\ \delta &= (-1)^{Np+N+1} * d* \end{aligned} \tag{3.7}$$

The Hodge Laplacian is defined in terms of the $d$ and $\delta$, as

$$\Delta := d\delta + \delta d \tag{3.8}$$

Note that on 0-forms, i.e. on $C^\infty(\mathcal{M})$, the Hodge Laplacian is equivalent to the Laplace-Beltrami operator. Indeed, for $f \in C^\infty(\mathcal{M})$ one has

$$\Delta f = \delta df + d\delta f = \delta df = -\mathbf{div}(\mathbf{grad}f) \tag{3.9}$$

where the second equality follows from the fact that $\delta$ annihilates 0-forms. The last step can be justified as follows. Consider $\mathbf{grad}f$. By its defining property it is the tangent

vector field dual to $df$ through the metric. It can be shown (see, for instance, [59]) that $\mathbf{div}(X) = *d * \tilde{X} = -\delta\tilde{X}$, where $\tilde{X}$ is the 1-form dual to $X$ through the metric.

In order to compute the leading symbol of the Hodge Laplacian, it is convenient to first compute the symbols of $d$ and $\delta$. This is an ideal setting to showcase the algebraic definition of the leading symbol introduced in Remark 2.2.2. In this endeavour, we follow [59]. Both operators considered are of first order, so it is sufficient to use a single $ad(f)$ map. Let $\omega \in \Omega^\bullet(\mathcal{M})$.

$$(ad(f)d)\,\omega = d(f\omega) - fd\omega = (df) \wedge \omega + fd\omega - fd\omega = df \wedge \omega \tag{3.10}$$

Recalling that $ad(f)$ only depends on $df$, one can replace $df$ by $\xi$ in the above. From that and Remark 2.2.2, one deduces that $\sigma_\xi(d) = e(\xi)$, where $e(\xi)$ is the left exterior product by $\xi$.

For the codifferential, one easily obtains

$$\sigma_\xi(\delta) = (-1)^{Np+N+1} * \sigma_\xi(d)* = (-1)^{Np+N+1} * e(\xi)* \tag{3.11}$$

as the Hodge $*$ commutes with multiplication by smooth functions and thus with $ad(f)$. This result can be significantly simplified. Let $\xi^* \in T_x(\mathcal{M})$ denote the metric dual of $\xi$ and let $i(\xi^*)$ be the interior derivative along $\xi^*$. It can be shown that $*e(\xi)* = -(-1)^{Np+N+1}i(\xi^*)$ [59]. Consequently, the principal symbol of $\delta$ is given by

$$\sigma_\xi(\delta) = -i(\xi^*) \tag{3.12}$$

Consider the Hodge-DeRham operator $d + \delta : \Omega^\bullet(\mathcal{M}) \to \Omega^\bullet(\mathcal{M})$. Its principal symbol is given by

$$\sigma_\xi(d + \delta) = \sigma_\xi(d) + \sigma_\xi(\delta) = e(\xi) - i(\xi^*) \tag{3.13}$$

As $d^2 = 0$ and $\delta^2 = 0$, $\Delta = (d + \delta)^2$. The rule of composition of principal symbols of Lemma 2.2.2 yields

$$\begin{aligned}\sigma_\xi(\Delta) &= \sigma_\xi((d+\delta)^2) = \sigma_\xi((d+\delta))^2 = (e(\xi) - i(\xi))^2 \\ &= -(e(\xi)i(\xi^*) + i(\xi^*)e(\xi)) = -\xi(\xi^*) = -\|\xi\|^2\end{aligned} \tag{3.14}$$

where the penultimate equality follows from the anti-Leibniz rule for interior derivatives. Recall that by Remark 2.2.2 the algebraic definition of the principal symbol differs by a

factor of $(-i)^m$ from the Fourier transform one. In this case, $m = 2$. The resulting factor of $-1$ ensures that the Hodge Laplacian is indeed of Laplace type.

### 3.1.3   Covariant Laplacian and Weitzenböck Remainder

Compared to the Laplace-Beltrami operator and the Hodge Laplacian, the covariant Laplacian is a rather crude creature. It is nonetheless important, as it allows to define Laplacians on any smooth vector bundle. In particular, the covariant Laplacian can be used to define a Laplace operator on tensors of any order. Let $E$ be a smooth vector bundle over the Riemannian manifold $(\mathcal{M}, g)$. Recall the definition of a linear connection on $E$.

**Definition 3.1.1** (Linear Connection). *A linear connection (or covariant derivative) on $E$ is a linear map*

$$\nabla : C^\infty(E) \to C^\infty(T^*\mathcal{M} \otimes E) \tag{3.15}$$

*such that, $\forall f \in C^\infty(\mathcal{M})$ and $\forall u \in C^\infty(E)$, one has*

$$\nabla(fu) = df \otimes u + f\nabla u \tag{3.16}$$

Suppose $E$ to be equipped with a Hermitian inner product $\langle \cdot, \cdot \rangle$. Together with $g$, this inner product induces a metric on $T^*\mathcal{M} \otimes E$. One can then find a formal adjoint $\nabla^*$ of the linear connection. We refer the reader to [59] for the details of this construction. One then can define the covariant Laplacian.

**Definition 3.1.2** (Covariant Laplacian). *Let $E, \nabla$ and $\nabla^*$ be as above. The covariant Laplacian on $E$ is defined to be*

$$\Delta_\nabla = \nabla^*\nabla \tag{3.17}$$

Notice the similarity of the above definition with that of the Laplace-Beltrami operator $\Delta = \delta d$. The covariant Laplacian is an operator of Laplace type. A proof of this fact can be found in [59]. Interestingly, any Laplace type operator can be expressed in terms of a covariant Laplacian.

**Theorem 3.1.1** (Weitzenböck Remainder). *Let $L : C^\infty(E) \to C^\infty(E)$ be an operator of Laplace type. Then, there exists a unique metric connection $\nabla$ on $E$ and an endomorphism $R(L)$ of $E$ such that*

$$L = \nabla^*\nabla + R(L) \tag{3.18}$$

*where $R(L)$ is said to be the Weitzenböck remainder.*

This can be used to simplify the study of arbitrary Laplace type operators, as is done in [8].

## 3.2 Geometry Reflected by Spectra

The spectra of Hodge Laplacians on compact Riemannian manifolds are known to contain a significant amount of geometric information. While generally insufficient to allow for a complete determination of a manifold, the information that one can extract from the spectrum (or spectra) using known techniques is still sufficient to establish some common features of isospectral manifolds. This section reviews some such results. A more complete review can be found in [5]. In subsequent chapters, those properties are used to prove certain positive results in inverse spectral geometry.

The earliest results on the encoding of geometric information in the spectra of operators come from the the study of the Laplace-Beltrami operator. Let $(\mathcal{M}, g)$ be a compact Riemannian manifold with boundary $\partial\mathcal{M}$. The expansions of the heat trace of the Laplace-Beltrami operator obeying Dirichlet or Neumann boundary conditions have similar forms, differing only by signs. In pairs of the form $\pm$ or $\mp$ the top sign will be the one used for Dirichlet boundary conditions and the bottom one the one used for Neumann conditions. Let $g_{\partial\mathcal{M}}$ denote the induced metric on the boundary. The first two expansion coefficients are

$$
\begin{aligned}
a_0 &= \frac{1}{(4\pi)^{N/2}} \int_{\mathcal{M}} dV_g = \frac{1}{(4\pi)^{N/2}} Vol(\mathcal{M}) \\
a_1 &= \mp\frac{1}{(4\pi)^{(N-1)/2}} \frac{1}{4} \int_{\partial\mathcal{M}} dV_{g_{\partial\mathcal{M}}} = \mp\frac{1}{(4\pi)^{(N-1)/2}} \frac{1}{4} Vol(\partial\mathcal{M})
\end{aligned}
\tag{3.19}
$$

After that, the coefficients become increasingly more and more complex integrals of combinatorial expressions involving various powers and contractions of the curvature tensor over $\mathcal{M}$ and $\partial\mathcal{M}$. If Laplacians on other bundles are used, the expressions also gain in complexity. For proofs of the expressions of higher order coefficients, we refer the reader to [8, 26].

The key message of Equation (3.19) is that the asymptotics of the heat trace encode the volume of the manifold and the length of the boundary. If $\partial\mathcal{M} = \varnothing$, the expressions for $a_0$ and $a_1$ still hold and one can thus still read the volume of the manifold from the heat trace asymptotics. This fact can be put in a particularly classical form.

**Theorem 3.2.1** (Weyl Estimate). *Let $\{\lambda_n\}_{n=1}^{\infty}$ be the spectrum of the Laplace-Beltrami operator on a compact $N$-dimensional Riemannian manifold $(\mathcal{M}, g)$ (with or without boundary). The eigenvalues obey the following estimate:*

$$
\lambda_k \sim \left( \frac{(2\pi)^N}{\beta(N)Vol(\mathcal{M})} \right)^{2/N} k^{2/N}
\tag{3.20}
$$

*where $\beta(N)$ is the volume of the unit ball in $\mathbb{R}^N$.*

The Weyl Estimate implies that all manifolds of a given dimension and volume have the same scaling properties of the eigenvalues. The intuition behind this is that eigenfunctions corresponding to high eigenvalues have very short wavelengths and thus are not sensitive the curvature or to the shape of the boundary. This result can be proven in two distinct ways. The original 1911 proof by Hermann Weyl was only valid for domains in $\mathbb{R}^2$. A generalization of it to bounded domains with smooth boundary $\mathcal{M} \in \mathbb{R}^N$ can be found in [4]. We will sketch this proof first.

*Sketch of Proof: Weyl Estimate 1.* The idea behind this proof is to formalize the intuition that eigenfunctions corresponding to high eigenvalues are not sensitive to the overall shape. This is done by splitting the manifold into subdomains on which the analysis is less involved. Consider a Dirichlet problem for the Laplace-Beltrami operator on $\mathcal{M}$, a bounded domain in $\mathbb{R}^N$ with smooth boundary. Consider a cubic grid of side $a$ in $\mathbb{R}^N$. For a given $a$, one identifies the cubes entirely contained in $\mathcal{M}$ and the cubes containing at least one point of $\mathcal{M}$. The union of the first set of cubes is denoted $C_I$ and the that of the second set $C_E$ where the subscripts stand for "interior" and "exterior", respectively. One clearly has $C_I \subset \mathcal{M} \subset C_E$. Then, an eigenvalue problem is considered on each separate cube of the lattice. The problems on the interior cubes are taken to satisfy Dirichlet boundary conditions while those on the exterior ones are taken to satisfy Neumann boundary conditions. Then, one collects the spectra of all cubes of $C_I$, *counting multiplicity* to form a nondecreasing sequence $\{\mu_n^I\}_{n=1}^\infty$. The same is done for the spectra of the cubes of $C_E$ to obtain a nondecreasing sequence $\{\mu_n^E\}_{n=1}^\infty$. The spectrum $\{\lambda_n\}_{n=1}^\infty$ of $\mathcal{M}$ can then be bounded as follows:

$$\mu_n^E \leq \lambda_n \leq \mu_n^I \tag{3.21}$$

This bound holds due to inequalities relating the spectra of Dirichlet and Neumann Laplacians on nested domains. Then, using the well known expressions for the eigenvalues of a cube of side $a$ and taking some appropriate limits the desired result is obtained. $\square$

An alternative proof of the Weyl Estimate uses the asymptotics of the heat kernel and a theorem about the relationship between the asymptotic behaviour of a measure on $(0, \infty)$ and its Laplace transform.

**Theorem 3.2.2** (Hardy, Littlewood, Karamata)**.** *Let $Q(\lambda)$ be a nondecreasing function such that the following Laplace transform converges for all $t \in (0, \infty)$*

$$\omega(t) = \int_0^\infty e^{-\lambda t} dQ(\lambda) \tag{3.22}$$

where the integral is taken in the Lebesgue-Stieltjes sense. Let $C, \rho$ be positive constants. The following two properties are equivalent:

$$\omega(t) \sim Ct^{-\rho} \quad as \ t \downarrow 0 \tag{3.23}$$

$$Q(\lambda) \sim \frac{C}{\Gamma(\rho+1)}\lambda^\rho \quad as \ \lambda \uparrow \infty \tag{3.24}$$

*Proof.* See [22]. It is worth the look. $\qquad\qquad\square$

**Remark 3.2.1.** *Implications that relate the measure to the transform ((3.24) $\implies$ (3.23)) are known as Abelian theorems, after Niels Henrik Abel . Implications that go the opposite way ((3.23) $\implies$ (3.24)) are known as Tauberian, after Alfred Tauber. In this context, Abelian is not the opposite of non-commutative, even though named after the same Abel [42].*

*Proof: Weyl Estimate 2.* This proof follows the strategy found in [18]. Notice that the expression of the heat trace for the Laplace-Beltrami operator can be seen as a Laplace transform with respect to the measure defined by $Q(\lambda)$, the function counting the number of eigenvalues lower or equal to $\lambda$

$$Tr(e^{-t\Delta}) = \sum_{k=1}^\infty e^{-\lambda_k t} = \int_0^\infty e^{-\lambda t} dQ(\lambda) \tag{3.25}$$

From the expansion of the heat trace of Theorem 2.4.2 and Equation (3.19). The heat trace obeys the following estimate:

$$Tr(e^{-t\Delta}) \sim \frac{1}{(4\pi)^{N/2}} Vol(\mathcal{M}) t^{-N/2} \tag{3.26}$$

The Tauberian part of theorem 3.2.2 then implies the following estimate

$$Q(\lambda) \sim \frac{1}{\Gamma(N/2+1)} \frac{1}{(4\pi)^{N/2}} Vol(\mathcal{M}) \lambda^{-N/2} = \frac{\beta(N)}{(2\pi)^N} Vol(\mathcal{M}) \lambda^{-N/2} \tag{3.27}$$

Since $Q(\lambda_k) = k$, this can be recast as

$$\lambda_k \sim \left( \frac{(2\pi)^N}{\beta(N)Vol(\mathcal{M})} \right)^{2/N} k^{2/N} \tag{3.28}$$

This is the desired result. $\qquad\square$

The expansion of the heat trace of the Hodge Laplacian is known to reflect some additional geometric information about the manifold. Let $\{\lambda_n^p\}_{n=1}^\infty$ denote the spectrum of the Hodge Laplacian acting on $p$-forms of a compact Riemannian manifold $(\mathcal{M}, g)$.

**Theorem 3.2.3.** *Let $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ be compact Riemannian manifolds without boundary of dimension $N \geq 2$ with Hodge Laplacian spectra $\{\lambda_n^p\}_{n=1}^\infty$ and $\{\mu_n^p\}_{n=1}^\infty$, respectively. If $\{\lambda_n^p\}_{n=1}^\infty = \{\mu_n^p\}_{n=1}^\infty$ for $p = 0, 1, 2$, the following hold:*

*(a) $(\mathcal{M}, g)$ is of constant curvature $c$ if and only if $(\mathcal{N}, h)$ is.*

*(b) $(\mathcal{M}, g)$ is of constant sectional curvature $c$ if and only if $(\mathcal{N}, h)$ is.*

*(c) $(\mathcal{M}, g)$ is Einstein if and only if $(\mathcal{N}, h)$ is.*

*Sketch of Proof.* The proof uses the explicit forms of the first few expansion coefficients of the heat trace for $p$-forms, which must be equal for isospectral manifolds. In fact, the isospectrality statement in the theorem can be relaxed to one of the equality of the first few coefficients. See [60] or [25] for details. $\qquad\square$

# Chapter 4

# Inverse Spectral Geometry

The goal of inverse spectral geometry is to extract as much information as possible from the spectrum of some differential operators defined on compact manifolds with or without boundary. It is known that the answer to Mark Kac's original question "Can One Hear the Shape of a Drum? " [45] is in general negative [27]. Nonetheless, there is a number of results in inverse spectral geometry one may consider positive. Some such results allow one to identify a shape from its spectrum in a suitably restricted class of shapes, others state that isospectral manifolds have some common geometric features. For instance, the original motivation for Kac's paper was the possibility to deduce the area and the boundary perimeter of a (sufficiently regular) planar domain from the spectrum of the Laplacian on scalar functions of this domain.

We begin by reviewing some negative and positive results in inverse spectral geometry. A third section surveys the topological properties of sets of isospectral manifolds. We close the chapter on a review of problems analogous to inverse spectral geometry, such as the determination of a Schrödinger potential from the spectrum of the quantum mechanical Hamiltonian.

## 4.1   Negative Results

Negative results in inverse spectral geometry are constructions of families of isospectral, non-isometric manifolds. In the simplest case, isospectral pairs are constructed. More involved cases include continuous families. This review is highly incomplete. The results presented here are those that we subjectively judged to be of historical importance and

often referred to in a way intimidating to a newcomer to the field. We thus claim to speak from experience. If one's only reliable method for solving partial differential equations is separation of variables, proving the isospectrality of sixteen-dimensional tori may seem quite tedious, considering how long it may take to solve PDEs in the relatively low dimensionality of space-time. Most daunting of all is the fact that Sunada's technique is often referred to as number-theoretic (by Sunada himself indeed [70], among others), which seems to imply that the newcomer needs to master an additional field of mathematics, on top of differential geometry and functional analysis, before even attempting to learn spectral geometry. Hopefully, the following shall dispel all such misgivings. A quite readable, more comprehensive list of negative results can be found in [29].

### 4.1.1  16-Dimensional Tori

In 1964, Milnor [55] constructed a pair of $16-$dimensional flat tori isospectral for Hodge Laplacians on $p-$forms. In [45], Kac cites this fact as reason to doubt the possibility of general success of the inverse spectral geometry program. Milnor's argument is an explicit construction that relies on a result about lattices in dimension sixteen, so no tedious separation of variables is needed. In fact, Milnor's original paper is only one page long.

A flat torus of dimension $N$ is defined as a Riemannian quotient manifold $\mathbb{R}^N/L$, where $L$ is a lattice (discrete additive subgroup of $\mathbb{R}^N$) of rank $N$. The dual lattice $L^*$ of $L$ consists of all $y \in \mathbb{R}^N$ such that $x \cdot y \in \mathbb{Z}$, for $x \in L$. All eigenfunctions of the Laplacian on scalar functions of $\mathbb{R}^N/L$ are given by $\psi_y(x) = \exp(2\pi x \cdot y)$ for $y \in L^*$. The corresponding eigenvalues have the form $\lambda_y = (2\pi)^2 y \cdot y$. Consequently, the spectrum of such a torus is uniquely determined by the function $N_L(r)$ counting the number of elements of $L^*$ in a ball of radius $r$ centered at the origin of $\mathbb{R}^N$. Milnor then invokes the fact that there exist two distinct self-dual lattices $L_1$ and $L_2$ in $\mathbb{R}^{16}$ such that $N_{L_1}(r) = N_{L_2}(r)$. This ensures the isospectrality of $\mathbb{R}^{16}/L_1$ and $\mathbb{R}^{16}/L_2$. The fact that the lattices are distinct (not equivalent by rotation) means that the tori obtained are non-isometric. Moreover, the isospectrality can be shown to persist for the Hodge Laplacian on $p-$forms.

### 4.1.2  Sunada Construction

In 1985, Sunada [70] introduced a powerful technique that allows one to construct pairs of isospectral manifolds. It relies upon factoring Riemannian manifolds with respect to subgroups of the isometry group. The treatment we present follows [29]. First, it is

necessary to introduce the notion of Riemannian covering. We refer the reader to [5] for further details.

**Definition 4.1.1** (Covering Space). *Let $\mathcal{M}, \mathcal{N}$ be differentiable manifolds. A map $p : \mathcal{M} \to \mathcal{N}$ is said to be a covering map if $\mathcal{N}$ is covered by open sets $U_i$ such that for each $U_i$, the set $p^{-1}(U_i)$ is a countable set of open subsets of $\mathcal{M}$ individually diffeomorphic to $U_i$. $\mathcal{M}$ is said to be a covering space of the base space $\mathcal{N}$.*

**Definition 4.1.2** (Deck Transformation). *A map $T : \mathcal{M} \to \mathcal{M}$ is said to be a deck transformation of a covering map $p : \mathcal{M} \to \mathcal{N}$ if it commutes with $p$. Deck transformations form a discrete subgroup of the full diffeomorphism group of $\mathcal{M}$.*

**Remark 4.1.1.** *The terminology in the above definitions can be visualized as follows. Imagine for simplicity that $\mathcal{N}$ is covered by a single open set $U$. The pre-image $p^{-1}(U)$ can be viewed as a stack of copies of $U$ on top of it. If $U$ is visualized as a rectangle, the stack of copies looks like a deck of cards, justifying the term deck transformation. Some authors prefer to view $U$ as a plate and $p^{-1}(U)$ as a stack of pancakes.*

**Definition 4.1.3** (Local Isometry). *A smooth map $b : \mathcal{M} \to \mathcal{N}$ between Riemannian manifolds $(\mathcal{M}, g)$ and $(\mathcal{N}, h)$ is said to be a local isometry if for every $x \in \mathcal{M}$, there is an open neighbourhood $U$ such that $b(U)$ is open in $\mathcal{N}$ and $b|_U : U \to b(U)$ is an isometry [44].*

**Definition 4.1.4** (Riemannian Covering). *Let $(\mathcal{N}, h)$ be a Riemannian manifold and let $p : \mathcal{M} \to \mathcal{N}$ be a covering map. $\mathcal{M}$ then inherits a pull-back metric $g = p^*h$. The pull-back is defined with respect to the diffeomorphism relating the $U_i$ and the discrete sets in the pre-image. The covering map $p$ is then a local isometry. Deck transformations of $p$ are local isometries on $(\mathcal{M}, g)$.*

**Definition 4.1.5** (Riemannian Quotient). *Let $(\mathcal{M}, g)$ be a Riemannian manifold and let $p : \mathcal{M} \to \mathcal{N}$ be a covering map. $\mathcal{N}$ then inherits a quotient metric $h = g/p$ by requiring that $p$ is everywhere a local isometry. Deck transformations of $p$ are local isometries on $(\mathcal{M}, g)$.*

**Remark 4.1.2.** *Given a discrete group $G$ acting on the left on $(\mathcal{M}, g)$ by isometries, one can define the Riemannian Quotient of $(\mathcal{M}, g)$ by this group. Let $\mathcal{N}$ be the manifold of orbits of $\mathcal{M}$ under the action of $G$ and let $p : \mathcal{M} \to \mathcal{N}$ be the projection onto the equivalence class. The projection $p$ is then a covering map and $\mathcal{N}$ inherits a quotient metric $h = g/p$. This construction is denoted $(\mathcal{N}, h) = (G \setminus \mathcal{M}, g)$.*

The Sunada construction relies on the observation that one can obtain the eigenfunctions of the Laplacian on the base space $(\mathcal{N}, h)$ from the eigenfunctions of the covering space $(\mathcal{M}, g)$. With that in mind, one factors $(\mathcal{M}, g)$ with respect to carefully chosen subgroups of its symmetry group to obtain isospectral non isometric manifolds. The appropriate criterion for that choice is given by the notion of almost conjugate subgroups. First, recall the more familiar notion of conjugate subgroups [72].

**Definition 4.1.6** (Conjugate Subgroups). *Two subgroups $\Gamma_1$ and $\Gamma_2$ of a group $G$ are conjugate if there exists $g \in G$ such that*

$$\Gamma_2 = g\Gamma_1 g^{-1} \tag{4.1}$$

**Definition 4.1.7** (Almost Conjugate Subgroups). *Let $\Gamma_1$ and $\Gamma_2$ be subgroups of a finite group $G$. $\Gamma_1$ is said to be almost conjugate to $\Gamma_2$ if each conjugacy class of $G$ intersects $\Gamma_1$ and $\Gamma_2$ in the same number of elements.*

As the terminology suggests, conjugate subgroups are also almost conjugate. The usefulness of the notion of almost conjugacy comes from the fact that it is equivalent to that of representation equivalence for finite groups. This property is used to prove the following theorem, which is a generalization of the one originally shown by Sunada [70].

**Theorem 4.1.1** (Sunada, Generalized). *Let $(\mathcal{M}, g)$ be a compact Riemannian manifold, with or without boundary, on which a finite group $G$ acts on the left by isometries. Let $\Gamma_1$ and $\Gamma_2$ be almost conjugate subgroups of $G$. If $\Gamma_1$ and $\Gamma_2$ act freely, the Riemannian quotients $(\Gamma_1 \setminus \mathcal{M}, g)$ and $(\Gamma_2 \setminus \mathcal{M}, g)$ are isospectral for all natural Laplace type operators.*

*Sketch of Proof.* Let $P$ be a natural Laplace type operator on some smooth vector bundle of $(\mathcal{M}, g)$. $P$ then *commutes with isometries*. In particular, $P$ commutes with the action of $G$. Consequently, the action of $G$ can be seen as mapping eigenspaces of $P$ into themselves (but not necessarily eigenfunctions into themselves). The eigenfunctions left invariant by the action of $\Gamma_1$ (respectively $\Gamma_2$) can be shown to be the eigenfunctions of $(\Gamma_1 \setminus \mathcal{M}, g)$ (respectively, $(\Gamma_2 \setminus \mathcal{M}, g)$). If $\mathcal{M}$ has boundaries, the boundary conditions imposed on the Riemannian quotients must be consistent with the ones imposed on $\mathcal{M}$. Thus, the spectra of $(\Gamma_1 \setminus \mathcal{M}, g)$ and $(\Gamma_2 \setminus \mathcal{M}, g)$ contain the same eigenvalues, not counting multiplicities. In order to show isospectrality, it remains to ascertain that the eigenspaces corresponding to identical eigenvalues have the same dimension. This conclusion can be reached from representation theoretical techniques applied to representations of $(\Gamma_1 \setminus \mathcal{M}, g)$ and $(\Gamma_2 \setminus \mathcal{M}, g)$ on finite dimensional vector spaces (eigenspaces of $P$). The almost conjugacy hypothesis is crucial for this last step, as it is equivalent to representation equivalence of the subgroups [29]. ☐

**Remark 4.1.3.** *($\Gamma_1 \setminus \mathcal{M}, g$) and ($\Gamma_2 \setminus \mathcal{M}, g$) are isometric if and only if $\Gamma_1$ and $\Gamma_2$ are conjugate subgroups of the* full *isometry group of $(\mathcal{M}, g)$. In practice, it is only possible to get a subgroup $G$ of the full isometry group such that the conditions of theorem 4.1.1 apply. Thus, in order to make sure that some specific ($\Gamma_1 \setminus \mathcal{M}, g$) and ($\Gamma_2 \setminus \mathcal{M}, g$) constructed by the Sunada method are indeed a counterexample to inverse spectral geometry, one must use some external method to check that they are not accidentally isometric.*

For examples of explicit constructions of counterexamples to inverse spectral geometry by the Sunada theorem and extensions of this technique, see [29] and the references therein. For a survey of various proofs and applications of Sunada's construction, see [11]. Note that isospectral pairs constructed by the Sunada method are locally isometric, even if they are not so globally.

### 4.1.3 Other Negative Results

**Locally Non-Isometric Isospectral Manifolds**

As mentioned above, the Sunada construction results in pairs of locally isometric isospectral manifolds. It is thus of immediate interest to attempt to construct pairs of locally non-isometric isospectral manifolds. The first examples of such manifolds were obtained by Szabo [71]. Szabo's proof relies on an explicit computations of the eigenvalues of Laplace-Beltrami operator on manifolds of the form $B_\delta \times T^3$, where $B_\delta$ is a ball in $\mathbb{R}^N$ and $T^3$ is a 3-torus, equipped with carefully chosen locally non-isometric metrics. Isospectrality holds for both Dirichlet and Neumann boundary conditions. In addition, the boundaries of those manifolds are isospectral but locally non-isometric.

**Continuous Isospectral Deformations**

The isospectral non-isometric manifolds produced by the Sunada technique generically come in discrete pairs. An interesting question to ask is whether one can construct continuous families of isospectral manifolds. This can be achieved by a number of constructions. For a review, see [29].

The first construction we explore is quite similar to Sunada's. The idea is to use Lie groups as both the manifold on which the Laplacian eigenvalue problem is defined and the symmetry group. Any compact Lie group $\mathcal{G}$ is naturally equipped with a metric $g$ invariant under left action by $\mathcal{G}$. One then selects a continuous family of discrete subgroups of $\mathcal{G}$ that

are representation equivalent and satisfy a series of other technical assumptions. Similarly to Sunada's construction, for any two such subgroups, the Riemannian quotient of $(\mathcal{G}, g)$ by those subgroups yields isospectral non-isometric manifolds. The continuity of the family of subgroups ensures the continuity of the family of isospectral non-isometric manifolds. The proof of this result can be found in [28]. Note that it predates Sunada's.

**Riemannian Submersions**

Recall that a submersion is a smooth map $\pi : \mathcal{M} \to \mathcal{N}$ between manifolds such that the associated push-forward $\pi_*$ is a surjective map of tangent spaces. A submersion is said to be Riemannian if this push-forward is an isometry between the tangent space of $(\mathcal{N}, h)$ and the orthogonal complement of the kernel of $\pi$ [5]. Under suitable additional hypotheses, one can relate the Laplace-Beltrami operator on $(\mathcal{M}, g)$ to the one on $(\mathcal{N}, h)$ as follows

$$\pi^*\Delta_\mathcal{N}(f) = \Delta_\mathcal{M}\pi^*(f) \ , \forall f \in C^\infty(\mathcal{N}) \tag{4.2}$$

Clearly, one then has that the spectrum of $\Delta_\mathcal{N}$ is equal to the spectrum of $\Delta_\mathcal{M}$ acting on $\pi^*C^\infty(\mathcal{N})$. Then, one uses this fact together with some additional special circumstances to obtain a pair of isospectral non-isometric manifolds. This can be extended to be valid for the Hodge Laplacian acting on exterior forms of any order [24]. Riemannian submersions can be used to obtain both continuous families of isospectral manifolds and pairs of locally non-isometric isospectral manifolds. Szabo's construction can be recast in those terms. See [29] and the references therein.

## 4.2  Positive Results

This section's aim is to review some positive results in inverse spectral geometry. Such results can be broadly divided in three classes. The strongest type of result ensures that some given geometry can be unambiguously identified amongst a set of similar objects by the spectra of some number of (pseudo-)differential operators defined on it. This is known as *spectral uniqueness*. The ultimate aim of inverse spectral geometry is to establish spectral uniqueness for all possible shapes. As it stands, uniqueness results are restricted to highly symmetric cases and are few. The conditions for spectral uniqueness can be weakened by only requiring that the spectrum (or spectra) of the geometry is unique amongst a set of shapes that are nearby in a suitable sense. This is known as *local uniqueness*. An even weaker property a geometry may have is known as *spectral rigidity*, which means that this

particular shape admits no nontrivial continuous isospectral deformation. It is clear that spectrally unique shapes are also locally unique and that locally unique shapes are also spectrally rigid. This terminology is the one adopted in a 2011 review of positive results in inverse spectral theory by Datchev and Hezari [17]. The discussion below follows that survey and the references therein.

## 4.2.1 Spectral Uniqueness

The following result establishes the spectral uniqueness of ball-shaped domains amongst compact subsets of $\mathbb{R}^N$ with compact boundary.

**Theorem 4.2.1** (Uniqueness of Balls). *Let $U \subset \mathbb{R}^N$ be an open set with smooth boundary such that the spectrum of the Laplace-Beltrami operator with Dirichlet or Neumann boundary conditions on it coincides with the one for the unit ball in $B_1 \subset \mathbb{R}^N$. Then $U$ is isometric to $B_1$.*

*Proof.* The proof relies on the first two coefficients in the heat trace expansion, which, for a compact manifold $U$ with smooth boundary $\partial U$, provide one with the volume of both $U$ and $\partial U$ (Equation (3.19)). Since, the heat trace is a spectral invariant, $U$ and $\partial U$ have the same volumes as $B_1$ and $\partial B_1$, respectively. The isoperimetric inequality states that the ratio $Vol(\partial \mathcal{M})/Vol(\mathcal{M})$ is minimized by balls. Thus, $U$ is isometric to the unit ball. $\square$

**Remark 4.2.1.** *The isoperimetric inequality can be proven using the methods of geometric measure theory. See, for instance, [21].*

Recall that the $N-$dimensional sphere $S^N$ is defined as the set of points of unit Euclidean norm in $\mathbb{R}^{N+1}$. The standard metric is inherited from $\mathbb{R}^{N+1}$ and is sometimes referred to as the round metric. It is tempting to believe that a similar uniqueness result holds for standard $N-$dimensional spheres when considered amongst all compact Riemannian manifolds without boundary. This is however not exactly so.

**Theorem 4.2.2** (Uniqueness of Spheres, $N \leq 6$). *Let $(\mathcal{M}, g)$ be a compact Riemannian $N-$dimensional manifold without boundary. Let the spectrum of the Laplace-Beltrami operator on $(\mathcal{M}, g)$ coincide with the spectrum of the Laplace-Beltrami operator on the standard sphere $S^N$ equipped with the round metric. If $N \leq 6$, this implies that $(\mathcal{M}, g)$ is isometric to the standard $N-$sphere.*

*Proof.* The proof uses the first four coefficients of the heat kernel expansion. See [73]. $\square$

In higher dimensions, the above result is not known to hold. However, if one is allowed to use the spectra of the Hodge Laplacians on forms, then it generalizes to any dimension.

**Theorem 4.2.3** (Uniqueness of Spheres). *Let $(\mathcal{M}, g)$ be a compact Riemannian manifold without boundary of dimension $N$. Let the spectrum of the Hodge Laplacians on $0, 1$ and $2$-forms on $(\mathcal{M}, g)$ coincide with those on the standard sphere $S^N$ equipped with the round metric. Then $(\mathcal{M}, g)$ is isometric to the standard sphere.*

*Proof.* Note that if $N = 1$, 2-forms can not be defined. The theorem still holds by virtue of Theorem 4.2.2. Otherwise, it holds due to Theorem 3.2.3, which guarantees that $(\mathcal{M}, g)$ is a of constant sectional curvature, which makes it a space form of the same curvature as the standard sphere, which is also a space form. Space forms are unique up to isometry [5]. This completes the proof. $\qquad\square$

A step away from spheres lie surfaces of revolution in $\mathbb{R}^{N+1}$ equipped with the induced metric. If the surface has boundaries, boundary conditions are required to also satisfy rotational symmetry. By separation of variables, the spectrum $\{\lambda_{mk}\}$ of the Laplace-Beltrami operator on such a surface is determined by the spectrum $\{m\}$ of the generator of rotations $\partial/\partial\phi$ and the spectrum of a Sturm-Liouville operator on a line segment. Under some non-degeneracy conditions, knowledge of the joint spectrum $(\sqrt{\lambda_{mk}}, m)$ is equivalent to the knowledge of the surface of revolution. This is proven in [35] using semiclassical approximations of the Schrödinger equation. For axially symmetric metrics on $S^2$ satisfying some additional non-degeneracy conditions, the spectrum of the Laplace-Beltrami operator has been shown to be sufficient in [76]. The non-degeneracy conditions ensure that the spectrum of the generator of rotations can be obtained from the Laplacian spectrum.

Another very interesting spectral uniqueness result [77] is known for a particular class of symmetric domains in $\mathbb{R}^2$. We can only sketch the proof of this result, as it requires techniques from the theory of dynamical billiards, which are beyond the scope of our presentation. Key to the proof is the notion of wave trace $Tr(\cos(t\sqrt{\Delta}))$, an object analogous to the heat trace, but for the wave equation. See [61] for a precise definiton. Like the heat trace, the wave trace is determined from the spectrum. Unlike heat, waves do not dissipate, but oscillate forever, which renders an expansion about $t = 0$ meaningless. Instead, one uses expansions about lengths of periodic orbits of an idealized billiard ball bouncing in the domain. Note that there is no difference between length and duration of an orbit, as the speed of the abstract waves is set to one. This expansion yields coefficients analogous to those of of the heat equation. The result in question relies upon particular hypotheses

about the shape of the boundary of a domain in $\mathbb{R}^2$ and the set of lengths of periodic billiard orbits, counting multiplicity. This set of lengths is known as the length spectrum. A periodic orbit length that appears only once in the spectrum is said to be non-degenerate.

Let $(x, y)$ be the standard cartesian coordinates in $\mathbb{R}^2$. Let $a > 0$ and let $f(x)$ be an analytic function such that $f(a) = f(-a) = 0$ and $f(x) \neq 0$ for $x \in (-a, a)$. Let $\Omega \subset \mathbb{R}^2$ be a domain bounded by $f(x)$, that is, the set of points such that $-a \leq x \leq a$ and $-|f(x)| \leq y \leq |f(x)|$. Suppose $\Omega$ has a non-degenerate vertical bouncing ball orbit of length $T$ such that orbits of length $2Tr$ for $r \in \mathbb{N}$ are also non-degenerate. Finally suppose that the orbit in question does not have critical points of $\pm f(x)$ for endpoints.

**Theorem 4.2.4.** *Let $\Omega$ and $\Omega'$ be as above. Suppose $\Omega$ and $\Omega'$ are isospectral with respect to the Laplace-Beltrami operator with either Dirichlet or Neumann boundary conditions. Then $\Omega$ is isometric to $\Omega'$*

*Proof.* See [77]. The analyticity of $f(x)$ allows to describe it in terms of its Taylor coefficients. The coefficients are obtained via perturbative methods from the heat trace. □

A result similar to the above can be obtained for domains in $\mathbb{R}^N$ bounded by analytic functions with reflexion symmetry about all axes and satisfying a hypothesis about bouncing ball orbits similar to the above [40]. To our knowledge, this is the most powerful high-dimensional spectral uniqueness result.

## 4.2.2   Local Uniqueness and Spectral Rigidity

Local spectral uniqueness and rigidity results are significantly weaker than global uniqueness results, even when the most interesting uniqueness results are only valid for classes of highly symmetric manifolds. The review of this section is quite incomplete, as some manifolds exhibiting local uniqueness or rigidity behaviour are very special cases. For further results, see [17].

It is known that ellipses are spectrally rigid amongst smooth domains in $\mathbb{R}^2$ that share the symmetries of the ellipse. The proof can be found in [39] and relies on wave trace techniques.

In the previous section, it was said that it is not known if standard spheres of dimension greater than six have a unique Laplace-Beltrami spectrum. It is however known that they are locally spectrally unique. This is shown in [74] by analyzing the behaviour of the first four heat trace expansion coefficients for a metric of nearly constant curvature 1. As it

turns out it is not possible for a manifold to be simultaneously isospectral to the standard $N$-sphere and have a curvature close to 1 in a suitable sense without being isometric to the sphere. A similar result holds for flat compact manifolds without boundary [50]. The proof relies on the coefficient $a_2$ of the heat trace expansion. This coefficient is zero for flat manifolds and, being a spectral invariant, for all manifolds isospectral to flat manifolds. Similarly to the sphere case above, it is impossible for a manifold to be simultaneously close and isospectral to a flat manifold without being isometric to it.

Compact negatively curved manifolds form another rather wide class of manifolds with rigidity and local uniqueness properties. A manifold is said to have a simple length spectrum if the ratio of the lengths of any two distinct closed geodesics is irrational. This condition is generic. In [16] it is shown that compact negatively curved manifolds with simple length spectrum are spectrally rigid. The proof involves techniques from the theory of dynamical systems, the fact that the spectrum of the Laplace-Beltrami operator determines the length spectrum on such manifolds and a property of integrals of smooth functions over closed geodesics. See [33] for an earlier proof for the two dimensional case. With additional hypotheses on the length spectrum, this result can be extended to one of local uniqueness [68].

## 4.3   Compactness of Isospectral Sets

Having established the existence of non-isospectral isometric manifolds and thus of counterexamples to inverse spectral geometry, it is of immediate interest to determine how large are the sets of manifolds with a given spectrum. In other words, it is of interest to determine whether, given an appropriate topology, isospectral sets are compact in some space of admissible geometries.

Such a compactness result is known for the Laplace-Beltrami operator on compact domains $U \subset \mathbb{R}^2$ with smooth boundary $\partial U$. The original proof is due to Melrose and can be found in [53]. The topology used there is the $C^\infty$ topology on the boundary curvature function $\kappa(s)$, where $s$ is the arclength of the boundary. This topology is well defined for the set of domains of a fixed boundary length $L$. Melrose shows that isospectral sets, elements of which necessarily have equal boundary lengths $L$, are compact subsets of the set of all domains with that particular boundary length. The proof uses perturbative expansions of the heat kernel coefficients of odd order. A review of alternative proofs of this result can be found in [36].

For the Laplace-Beltrami operator on compact Riemannian manifolds without boundary, a more powerful result has been proven by Zhou [78]. For a given compact manifold

without boundary, sets of isospectral metrics with uniformly bounded sectional curvatures are compact in the $C^\infty$ topology. The proof uses the coefficients of the heat kernel expansion to bound the Sobolev norms of the isospectral metrics. For a review of other similar results, see [17].

## 4.4 Analogous Problems

Inverse spectral geometry can be viewed as part of a larger class of inverse problems with different physical interpretations, all of which involve the recovery of what is traditionally viewed as the conditions of a problem from the spectra of some suitable operators. In this section, we present some such problems which can be used as somewhat simpler analogues of inverse spectral geometry.

### 4.4.1 Hearing Boundary Conditions

The results in inverse spectral geometry presented in earlier sections of this chapter pertain to a particular brand of spectral geometry. The data one wished to recover from spectra were of differential geometric nature, as the ultimate goal was to recover the metric and the shape of the boundary. Instead, one can ask if, on a given Riemannian manifold $(\mathcal{M}, g)$ with boundary $\partial\mathcal{M}$, it is possible to recover the boundary conditions from the spectrum of some operator (typically a Laplacian) obeying those conditions. In essence, the goal is to use the spectrum to determine a particular function of $\partial\mathcal{M}$ that defines boundary conditions.

A negative result for discontinuously variable boundary conditions has been obtained in [43]. There, the boundary $\partial\Omega$ of a bounded domain $\Omega \subset \mathbb{R}^2$ is decomposed as $\overline{\partial\Omega_1 \cup \partial\Omega_2}$ where each $\partial\Omega_i$ is a union of open segments of the boundary $\partial\Omega$ and $\partial\Omega_1 \cap \partial\Omega_2 = \emptyset$. Note that $\partial\Omega_1$ and $\partial\Omega_2$ can be arbitrarily close to each other. Moreover, it is supposed that there is no nontrivial isometry of $\mathbb{R}^2$ that interchanges $\partial\Omega_1$ with $\partial\Omega_2$. One then imposes Dirichlet boundary conditions on one $\partial\Omega_i$ and Neumann boundary conditions on the other. Generically, the conditions would thus change discontinuously from Dirichlet to Neumann along the boundary. This is known as *Zaremba boundary conditions*, after Stanislas Zaremba who introduced them in [75]. It is possible to show that the spectrum of the Laplace-Beltrami operator cannot always distinguish which condition is imposed on which part of the boundary. Using Sunada's technique, one can construct non-trivial partitions of the boundary of certain planar domains, including a half-disk, such that

the spectrum of the Laplace-Beltrami operator is invariant under the interchange of the boundary conditions imposed on the $\partial \Omega_i$.

On the positive side, one can hear a particular kind of boundary condition on the ellipse. Let $\Omega$ be an ellipse with a boundary $\partial \Omega$ given by $x^2/a + y^2/b = 1$ for $a > b > 0$. Let $K$ be a smooth function of $\partial \Omega$ such that $K$ is even in $x$ and $y$. In other words, $K$ satisfies the reflection symmetries of the ellipse. Denoting directional derivatives with respect to the normal to the boundary by $\partial/\partial n$, and letting $u \in C^{\infty}(\Omega)$, the following boundary condition is defined:

$$\frac{\partial u}{\partial n} = Ku \tag{4.3}$$

In [34] it is shown that the spectrum of the Laplace-Beltrami operator with the above boundary condition uniquely determines $K$. The proof uses techniques from the theory of dynamical billiards.

## 4.4.2 Spectral Graphs

A very natural extension of inverse spectral geometry is the study of the shape of graphs through the spectrum of the *graph Laplacian* and the heat equation on graphs. This straightforward adaptation of spectral geometry to a discrete setting is called *spectral graph theory*. An introduction to the subject can be found in [14]. We follow that text to sketch the definition of graph Laplacian.

Let $\mathcal{G}$ be a graph with $N$ vertices. Let $d_i$ denote the degree of the node $i$. We say $i \sim j$ if the vertices $i$ and $j$ are adjacent. The (non-normalized) graph Laplacian is defined to be an $N$ by $N$ matrix $L$ with the following matrix elements:

$$L_{ij} = \begin{cases} d_i & \text{if } i = j \\ -1 & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

It can be convenient to normalize the above expression to make all nonzero diagonal elements equal to 1. This is not necessary for our discussion. The graph Laplacian acts by matrix multiplication on functions $f : \mathcal{G} \to \mathbb{R}, \mathbb{C}$, which can be viewed as $N$-dimensional vectors. Similarly to the geometric Laplacian, the sign convention is such that the eigenvalues of the graph Laplacian are real and non-negative. The physical meaning of the graph

Laplacian is best understood by studying the heat equation on it. Let $t \in (0, \infty)$ be a time parameter. The heat equation on a graph is defined to be

$$\frac{\partial u}{\partial t} = -Lu \qquad (4.5)$$

Select a node $i$ and consider the time derivative of the temperature at this node. As $u(i)$ increases, $\frac{\partial u(i)}{\partial t}$ decreases, indicating that heat wants to flow from hot to cold nodes. Similarly, as the temperature $u(j)$ of a node $j$ adjacent to $i$ increases, $\frac{\partial u(i)}{\partial t}$ increases. Notice that the factor $d_i$ counterbalances the number of nodes adjacent to $i$ so that if $u(i) = u(j)$ for all $j \sim i$, $\frac{\partial u(i)}{\partial t} = 0$. This can be viewed as the heat flow along a structure made of metal rods built in the shape of the graph $\mathcal{G}$, as long as one only cares about the temperatures at the nodes and abstracts out the heat transfer along the rods. Notice that in the above all adjacent nodes are equally important. This can be seen as all rods being of equal length. A more general setting is that of *weighted graphs*. Let $w(i, j)$ be a weight function satisfying

$$
\begin{aligned}
w(i, j) &\geq 0 \\
w(i, j) &= w(j, i) \\
w(i, j) &\neq 0 \iff i \sim j
\end{aligned}
\qquad (4.6)
$$

The definition of the degree of the node $i$ is modified to be $d_i = \sum_j w(i, j)$. The (non-normalized) graph Laplacian for the weighted graph is

$$
L_{ij} = \begin{cases} d_i - w(i, i) & \text{if } i = j \\ -w(i, j) & \text{if } i \sim j \\ 0 & \text{otherwise} \end{cases}
\qquad (4.7)
$$

Clearly, the unweighted graph Laplacian is a special case of the weighted one.

Spectral graph theory is analogous to spectral geometry as it tries to relate the connectivity properties of the graph and the weights of its edges to the spectrum of the graph Laplacian. The spectrum of the graph Laplacian is that of its matrix. Notice that it always admits a constant eigenfunction with eigenvalue zero. The multiplicity of the zero eigenvalue gives the number of components of the graph. Other than that, no information can be gathered from it. On the other hand, the smallest nonzero eigenvalue is known to contain a great deal of information about the connectivity properties of the graph. Note that for graphs of finite size, one can not study the asymptotic properties of the spectrum,

as there are no asymptotics to speak of. This is why attention is directed towards the lowest nontrivial eigenvalues. See [56] for a survey of basic results.

It is known that one can construct isospectral graphs with different geometric properties via an adaptation of the Sunada method [11]. Inverse spectral geometry of graphs thus has difficulties analogous to those of inverse spectral geometry of manifolds.

It is tempting to attempt to use graph Laplacians to approximate the Laplace-Beltrami operator on some manifold. Indeed, intuitively, one expects that by discretizing the manifold as a graph and taking the discretization to be finer and finer, one obtains a better and better approximation of the Laplace-Beltrami operator on that manifold. It is however not quite so. In order to ensure convergence to the Laplace-Beltrami operator, one has to use a correct normalization of the Laplacian. See [37] for details.

### 4.4.3 Schrödinger Potentials and Scattering

Perhaps the most physically relevant analogue to inverse spectral geometry is the determination of a quantum mechanical potential $V$ from the spectrum of the associated Schrödinger operator $S$. In $\mathbb{R}^N$ and in a geometrized system of units, this takes the form

$$S = \Delta + V \tag{4.8}$$

where $\Delta$ is the non-negative Laplace-Beltrami operator. Some potentials, such as the one for the quantum harmonic oscillator, naturally result in a discrete spectrum. Others may require the problem to be defined on a bounded domain $\Omega \subset \mathbb{R}^N$, supplied with appropriate boundary conditions. A related class of problems, known as potential scattering problems, study how Schrödinger waves coming from infinity interact with a compactly supported potential $V$ and then return to infinity. The operators involved have continuous spectrum on $[0, \infty)$ and possibly a finite number of negative eigenvalues. Thus, strictly speaking, the spectrum of $S$ contains very little geometric information. Instead, one uses the resonances associated with $S$. Let $(S - z^2)^{-1}$ be the resolvent of $S$. Modulo some technicalities with domains of operators and meromorphic continuations, resonances are defined to be the poles of the resolvent. They form a countable set. One can then attempt to recover $V$ from the set of resonances of $S$ just as if it was a spectrum. Scattering problems may also be defined in the absence of potentials but in the presence of obstacles. Obstacles are taken to be bounded subsets $O \subset \mathbb{R}^N$ with smooth boundary. The resonance problem is then defined on $\Omega = \mathbb{R}^N \backslash \overline{O}$ with some appropriate boundary conditions, usually Dirichlet

or Neumann. The goal is then to determine the shape of $O$ from the resonances. A review of positive results in potential identification and scattering can be found in [17].

A particularly curious negative result in the determination of a Schrödinger potential from its spectrum has been obtained in [52]. There, it is shown that there exists an infinite dimensional family of smooth potentials *isospectral to the quantum harmonic oscillator on the real line*. The proof relies on an explicit exploration of the geometry of this family via an isospectral flow. The resulting potentials look like $x^2$ for sufficiently large $x$ and have nontrivial structure near $x = 0$.

An equally curious positive result exists for the determination of potentials on the interval $[0, 1]$ with Dirichlet boundary conditions. In [62] it is shown that one can reconstruct the potential from three spectra, the spectrum of the full problem on $[0, 1]$, the spectrum of the problem restricted to $[0, 1/2]$ and the spectrum of the problem restricted to $[1/2, 1]$. In all cases, Dirichlet boundary conditions are imposed on the ends of the intervals. This reconstruction is not possible in full generality, as it requires for the three spectra to be pairwise disjoint. The result has later been extended to be valid with subintervals $[0, a]$ and $[a, 1]$, for $0 < a < 1$, as long as the three spectra are pairwise disjoint [23]. In all cases, proofs involve techniques from the theory of Sturm-Liouville operators.

# Chapter 5

# Spectral Triples

The theory of spectral triples originates in the field of non-commutative geometry, where it provides an alternative, coordinate free, description of manifolds. This subject is only tangentially related to inverse spectral geometry as formulated in the earlier chapters. Indeed, the description of geometry used in this chapter is based on more than just a spectrum. We choose to discuss it nonetheless as we believe it could be a valuable tool for inverse spectral geometry. Moreover, the theory of spectral triples is sometimes referred to as spectral geometry, causing confusion in the uninitiated.

A spectral triple is composed of three elements: a von Neumann algebra, a separable infinite-dimensional Hilbert space $\mathcal{H}$ and an operator on that space. If the algebra is taken to be the algebra of smooth functions $C^\infty(\mathcal{M})$ on a compact manifold $\mathcal{M}$ without boundary and the operator is taken to be a Dirac operator $D$ on that manifold, the Riemannian manifold $(\mathcal{M}, g)$ can be reconstructed from the knowledge of the *spectral triple*

$$(C^\infty(\mathcal{M}), \mathcal{H}, D) \tag{5.1}$$

Since all infinite-dimensional separable Hilbert spaces are isomorphic, $D$ can simply be taken to be the spectrum of the Dirac operator. Since the (geometric) Dirac operator is elliptic [25, 54, 59], one can then just choose to describe $\mathcal{H}$ in its eigenbasis without loss of generality. This is the reason why this technique is sometimes referred to as spectral geometry. The key difference between this approach and the one used in all the other chapters is the a priori knowledge of the algebra of functions, which will turn out to be a most powerful hypothesis. This chapter is a very limited introduction to this subject. We refer the reader to [15] for a full discussion.

## 5.1 Motivation for Spectral Triples

Consider a particle obeying the laws of non-relativistic quantum mechanics moving in one dimension. By the Heisenberg uncertainty principle, one may not simultaneously resolve its position and momentum. The uncertainties $\Delta x$ in position and $\Delta p$ in momentum satisfy:

$$\Delta x \Delta p \geq \frac{1}{2}\hbar \tag{5.2}$$

It is well known that this is due to the fact that the position and momentum operators in quantum mechanics do not commute. This contrasts with the case of classical mechanics, where those operators commute and one can resolve the phase space to arbitrary precision. The phase space of classical mechanics is well described in terms of differentiable manifolds. Quantum mechanical phase space, on the other hand, looks like the classical one with the uncertainty relation put on top of it. One is thus inclined to ask if it is possible to mathematically describe a space that naturally has such an unsharpness to its points, with differential geometry as a limiting or special case. Non-commutative geometry offers such a description, at least to an extent. For our purposes however, only the special commutative case is interesting. One can say that we are doing commutative non-commutative geometry.

The main idea behind the construction we are about to outline is to recast the problem of measuring distances between points in a Riemannian manifold in a quantum mechanics-like manner. That is, the goal is to recast the usual geometric way to measure distances in terms of operators.

Consider the real line equipped with the standard metric. In the standard coordinate system, the distance between two points $p$ and $q$ is given by $d(p, q) = |p - q|$. In general, however, the distance is set to be the infimum of the length of all paths beginning at $p$ and ending at $q$. The construction of spectral triples is based upon the observation that instead of using an infimum to define distance, one may use a cleverly constructed supremum.

$$d(p,q) = \sup_{f \in C^1(\mathbb{R})} \{|f(p) - f(q)| : |\partial_x f| \leq 1 \ , \ \forall x \in \mathbb{R}\} \tag{5.3}$$

The above works as $f(x) = x \in C^1(\mathbb{R})$ saturates the inequality and gives precisely the usual notion of distance. It indeed is the supremum, as any function $g \in C^1(\mathbb{R})$ such that $|g(p) - g(q)| > |p - q|$ can not have its derivative bounded by one without violating the mean value theorem, which would of course result in a contradiction. In higher dimensions, an analogous result is obtained by bounding the gradient. This, and all similar results, will

be referred to as Connes' distance formula, after Alain Connes, the main architect of non-commutative geometry. Notice that this approach is closer in spirit to the one encountered in quantum mechanics, as points and paths are replaced by the evaluation of functions at points, which can be viewed as the action of an operator.

## 5.2 Formulation of Spectral Triples

Let $(\mathcal{M}, g)$ be a compact Riemannian manifold. In this section, we shall establish the equivalence between the standard description of $(\mathcal{M}, g)$ and its spectral triple $(C^\infty(\mathcal{M}), \mathcal{H}, D)$. The previous section touched upon distance measurement based on the bounding of a first order differential operator. The Dirac operator of this section will play an analogous role. The Hilbert space $\mathcal{H}$ will be that of spinors, so the manifold $(\mathcal{M}, g)$ must be able to admit a spin structure. There are topological limitations to this and we refer the reader to [44] jointly with [41] for the technical details. A more pressing concern is that we no longer know the differentiable manifold $\mathcal{M}$. In the above motivational discussion, we knew precisely what the points were and thus could evaluate functions at them. This does not seem to be the case when first examining a spectral triple. Our first goal is thus to show that knowledge of $C^\infty(\mathcal{M})$ is perfectly equivalent to that of $\mathcal{M}$.

Given $\mathcal{M}$, a differentiable manifold, one can construct the space of smooth functions on it by the usual methods of differential geometry. The opposite way is not as simple. Consider a unital (equipped with a multiplicative identity $\mathbb{1}$) $\mathbb{R}$-algebra $\mathcal{F}$. Let $\mathbf{Spec}_\mathbb{R}(\mathcal{F})$ denote the set of all unital (mapping $\mathbb{1}$ to 1) algebra homomorphisms of $\mathcal{F}$ into $\mathbb{R}$. This set is called the spectrum of the algebra. In some cases, one can assign a geometric meaning to $\mathbf{Spec}_\mathbb{R}(\mathcal{F})$. The following theorem does that for algebras of smooth functions on differentiable manifolds.

**Theorem 5.2.1.** *Let $\mathcal{M}$ be a differentiable manifold and let $C^\infty(\mathcal{M})$ be its algebra of real-valued smooth functions. Then,*

$$\mathcal{M} = \mathbf{Spec}_\mathbb{R}(C^\infty(\mathcal{M})) \tag{5.4}$$

*Sketch of Proof.* Let $x \in \mathcal{M}$. Notice that the evaluation of smooth functions at $x$ is a unital algebra homomorphism from $C^\infty(\mathcal{M})$ to $\mathbb{R}$. By abuse of notation, $x$ and the evaluation at $x$ will be denoted by the same symbol: $x(f) := f(x)$ for $f \in C^\infty(\mathbb{R})$. First, we are to show that this is the only possible kind of homomorphism or that there is a bijection between $\mathcal{M}$ and $\mathbf{Spec}_\mathbb{R}(C^\infty(\mathcal{M}))$. Then, we are to establish that $\mathbf{Spec}_\mathbb{R}(C^\infty(\mathcal{M}))$ can be equipped

with a natural atlas compatible with the one on $\mathcal{M}$. We will limit our proof to the first statement, only sketching the second. See [58] for full details.

Let $\theta : \mathcal{M} \to \mathbf{Spec}_{\mathbb{R}}(C^\infty(\mathcal{M}))$ map points $x$ of $\mathcal{M}$ into the evaluation at $x$ homomorphism. This map is clearly injective, since for any $x \neq y \in \mathcal{M}$, there exists a function $f \in C^\infty(\mathcal{M})$ such that $f(x) \neq f(y)$.

To show surjectivity, consider some unital homomorphism $p : C^\infty(\mathcal{M}) \to \mathbb{R}$ and let $f \in C^\infty(\mathcal{M})$ be a function with compact level sets. See [58] for a proof of the existence of such functions. Let $L = f^{-1}(\lambda)$ be such a level set. Suppose that there is no $x \in L$ such that $p$ corresponds to the evaluation at $x$ homomorphism. Then the following family of functions exists $\{f_x \in C^\infty(\mathcal{M}) : x \in L, f_x(x) \neq p(f_x)\}$. $L$ has the following open cover:

$$L = \bigcup_x U_x \ , \ U_x = \{q \in \mathcal{M}; f_x(q) \neq p(f_x)\} \tag{5.5}$$

Since $L$ is compact, there exists a finite subcover $U_{x_1}...U_{x_m}$. The following smooth function is nonzero everywhere on $\mathcal{M}$:

$$g = (f - \lambda)^2 + \sum_{k=1}^{m}(f_{x_k} - p(f(x_k)))^2 \tag{5.6}$$

Since $g$ is nowhere zero, $1/g$ exists and is smooth. Notice that $g$ was constructed so that $p(g) = 0$. We obtain the following contradiction:

$$1 = p\left(g \cdot \frac{1}{g}\right) = p(g)p\left(\frac{1}{g}\right) = 0 \tag{5.7}$$

Thus, $p$ has to be an evaluation at a point homomorphism. There thus exists a natural bijection between $\mathcal{M}$ and $\mathbf{Spec}_{\mathbb{R}}(C^\infty(\mathcal{M}))$. The topology on $\mathbf{Spec}_{\mathbb{R}}(C^\infty(\mathcal{M}))$ is chosen to be the weakest topology such that $x(f) = f(x)$ defines continuous functions. That topology is Haussdorff. Proving that one can recover the appropriate differentiable structure is quite lengthy, so we refer the reader to [58]. Intuitively, beforehand knowledge of $C^\infty(\mathcal{M})$ tells one which continuous real functions are smooth and which ones are not. $\qquad\square$

**Remark 5.2.1.** *The above proof can be used to generalize the notion of differential calculus on manifolds to various other geometric objects through the algebraic definition of derivatives and the identification between the points of the object and the spectrum of its algebra of functions. This places differential calculus entirely within the realm of commutative algebra. A physical motivation for this approach is that if $\mathcal{M}$ represents the space of*

*configurations of a physical system, $C^\infty(\mathcal{M})$ is the mathematical model for a laboratory, each function being an instrument. To read the measurement of an instrument, one applies an element of $\mathbf{Spec}_\mathbb{R}(C^\infty(\mathcal{M}))$ onto it. Defining calculus over the instruments rather than the configuration space corresponds to only studying what one can measure. This approach is explored in detail in [58]. The ultimate goal of the program outlined there is to obtain a better mathematical description of quantum field theory.*

Now that we are able to extract the differentiable manifold from the spectral triple, we can obtain the metric in a way analogous to the one used in the previous section. First, notice that bounding the derivative of a function $f$ of the real line is equivalent to bounding its commutator with the derivative operator:

$$[\partial_x, f]g = \partial_x(fg) - f\partial_x g = g(\partial_x f) \tag{5.8}$$

A completely analogous construction can be carried out for the Dirac operator $D$. It can be shown that the operator on $\mathcal{H}$ defined by the commutator of the Dirac operator with the multiplication by $f \in C^\infty(\mathcal{M})$ is bounded by the exterior derivative, and thus the gradient of $f$ [15]:

$$\|[D, f]\| \leq |df| \tag{5.9}$$

For $p, q \in \mathcal{M}$, Connes' distance formula becomes:

$$d(p, q) = \sup_{f \in C^\infty(\mathcal{M})} \{\|f(p) - f(q)\| : \|[D, f]\| \leq 1\} \tag{5.10}$$

The above formula indeed recovers the geodesic distance between the points $p$ and $q$ [15]. It can be shown that those mutual distances correspond to a Riemannian manifold unique up to diffeomorphism. Thus, the spectral triple $(C^\infty(\mathcal{M}), \mathcal{H}, D)$ uniquely encodes the Riemannian manifold $(\mathcal{M}, g)$. Moreover, it does so in a completely coordinate free way.

**Remark 5.2.2.** *Non-commutative geometry is obtained by replacing $C^\infty(\mathcal{M})$ by some suitable non-commutative algebra. The physical motivation for this program is to obtain the geometric structure of quantum field theory [15]. This approach is incompatible with the one outlined in Remark 5.2.1.*

**Remark 5.2.3.** *Connes' distance formula and spectral triples have been used to reformulate the Einstein-Hilbert action of General Relativity in terms of the spectrum of the Dirac operator [51].*

A few comments on the above are now in order. First, note that Laplacians can not be used in Connes' distance formula instead of the Dirac operator as they are not bounded by $|df|$ due to the fact that they are second order differential operators. Second, it is important to realize that a spectral triple contains much more information than the notation lets on. Indeed, Connes' distance formula requires one to know how $C^\infty(\mathcal{M})$ acts on $\mathcal{H}$. Knowledge of this action is thus provided by the spectral triple. Since $C^\infty(\mathcal{M})$ contains arbitrarily localized functions and the spectral triple supplies us with the action of $D$ on $\mathcal{H}$, we essentially also know the eigenfunctions of $D$. Consequently, one can not consider spectral triples as a solution to the inverse spectral geometry program. Nonetheless, the spectral triple may turn out to be a convenient way of thinking about Riemannian manifolds in an algebraic, coordinate free, way for the purposes of infinitesimal inverse spectral geometry. For instance, one could attempt to adapt the methods outlined in Remark 5.2.3.

Another possible application of Connes' distance formula could be a rigorous formulation of the ideas outlined in the introduction of this thesis. Recall that, following [47], we motivated inverse spectral geometry by heuristically replacing the measure of mutual distances by the sampling of quantum mechanical correlations. We conjecture that Connes' distance formula can be used to rigorously examine this procedure.

# Chapter 6

# Numerical Experiments in Infinitesimal ISG

If one wishes to apply inverse spectral geometry to any given problem, the results discussed in Chapter 4 leave a lot to be desired. The most general formulation of inverse spectral geometry is plagued by counterexamples, while the cases when reconstruction of shape from spectrum is possible require stringent symmetry properties and do not seem to be generalizable to different settings. It would thus be highly desirable to obtain a general method that does not require symmetry, would provide a way to find specific solutions and would, at least in spirit, generalize from one setting to the other.

The obvious obstacle to this program is that the map between shapes and spectra is highly nonlinear and thus difficult to handle. For the past few hundreds of years, the main technique to deal with hard nonlinear problems was to linearize them and try to solve them locally. For maps between shape and spectra, this corresponds to perturbation theory. Given that, it becomes tempting to try to construct inverse spectral geometry as an iteration of perturbative changes in shape, directed by small changes in spectrum in order to, overall, go from some initial shape to a shape with a desired spectrum.

Such an approach has recently been applied with success to a particular class of spectral graphs [1]. Due to the complexity of the problem, a numerical approach was used. We shall build upon that success by formulating a general framework for an *infinitesimal* version of inverse spectral geometry, which includes the method used in [1] as a special case. As we will need to differentiate the map between shape and spectrum with respect to changes in shape, we first state some theorems that ensure that such operations are allowed. Then, we formulate a general framework of infinitesimal inverse spectral geometry adapted for

71

numerical computations. This framework can be straightforwardly modified to suit any inverse spectral problem. We then discuss possible numerical implementations of this framework. Finally, we apply one such implementation to the inverse spectral geometry of a class of star-shaped domains in $\mathbb{R}^2$ and discuss the success and failure rates obtained.

## 6.1  Continuity of Eigenvalues

An infinitesimal approach to inverse spectral geometry requires that small changes in the shape of a manifold result in small changes in the spectrum. More accurately, the changes in question should be differentiable. In this section, we state some theorems that ensure that such a thing is possible. Moreover, we state results that guarantee that manifolds with degenerate eigenvalues are, in some sense, rare.

We first sketch a continuity of eigenvalues result for compact Riemannian manifolds without boundary. The full version of it can be found in [3]. Let $(\mathcal{M}, g)$ be an $N$-dimensional compact Riemannian manifold without boundary and let $\Delta_g$ be the Laplace-Beltrami operator associated with the metric $g$. In order to discuss the dependence of eigenvalues on the metric, one introduces a Fréchet space structure on $S(\mathcal{M})$, the space of smooth symmetric covariant 2-tensors, and a complete metric $\rho$ on the space $\mathcal{G}(\mathcal{M})$ of smooth Riemannian metrics on $\mathcal{M}$. It is then possible to show that the eigenvalues of the Laplace-Beltrami operator are continuous with respect to the metric in the following sense:

**Theorem 6.1.1** (Eigenvalue Continuity). *Let $g, h \in \mathcal{G}(\mathcal{M})$. Suppose $\delta > 0$ and $\rho(g, h) \leq \delta$. Then, for any $k \in \mathbb{N}\backslash\{0\}$,*

$$e^{-(N+1)\delta} \leq \frac{\lambda_k(g)}{\lambda_k(h)} \leq e^{(n+1)\delta} \tag{6.1}$$

*Proof.* See [3]. $\qquad\square$

**Remark 6.1.1.** *The above result has been generalized to one of uniform continuity [7]. This implies that if one imposes a cutoff $\Lambda$ on the spectrum of the Laplacian such that only the first $n$ eigenvalues such that $\lambda_i < \Lambda$ are considered, $n$ will not change under sufficiently small perturbations of the metric. That is, unless there is an eigenvalue such that $\lambda_i = \Lambda$, no eigenvalue will cross the cutoff $\Lambda$ for sufficiently small changes in the metric.*

For a parametrized change of the metric, the eigenvalues are differentiable in the following sense:

**Theorem 6.1.2** (Eigenvalue Differentiability). *For $g \in \mathcal{G}(\mathcal{M})$ and $h \in S(\mathcal{M})$, let $g(t) = g + th \in \mathcal{M}$, $|t| \le \varepsilon$. Let $\lambda$ be an eigenvalue of $\Delta_g$ with multiplicity $l$. Then, there exist parametrized eigenvalue-eigenfunction pairs $\{(\Lambda_i(t), \psi_i(t)\}_{i=1}^l$ such that*

*(a) For fixed $i$, $\Lambda_i(t)$ and $\psi_i(t)$ depend real-analytically on $|t| \le \varepsilon$*

*(b) For fixed $i$ and $t$, $\Delta_{g(t)}\psi_i(t) = \Lambda_i(t)\psi_i(t)$*

*(c) For fixed $i$, $\Lambda_i(0) = \lambda$*

*(d) For fixed $t$, $\{\psi_i(t)\}_{i=1}^l$ is orthonormal with respect to the $L_2$ inner product induced by $g(t)$.*

*Proof.* See [3]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

The following result guarantees that manifolds with degenerate eigenvalues are rare. Recall that a residual set is the complement of a meagre set, which itself is a countable union of nowhere dense sets.

**Theorem 6.1.3** (Uhlenbeck). *Let $\mathcal{M}$ be a compact manifold of dimension $N \ge 2$. Let*

$$\mathcal{S} = \{g \in \mathcal{G}(\mathcal{M}) : \Delta_g \text{ has no degenerate eigenvalues}\} \qquad (6.2)$$

*Then $S$ is a residual set in $(\mathcal{G}(\mathcal{M}), \rho)$.*

*Proof.* See [3]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

For bounded domains in $\mathbb{R}^N$, the corresponding theorems take a more complicated form, since the shape is no longer described by a function of the interior of the domain, but rather by the shape of the boundary. A general notion of continuity and derivatives is thus more complicated to obtain as one instead needs to obtain a notion of continuity of the boundary shapes. A similar problem arises in continuum mechanics where one studies the deformation of a chunk of elastic material whose shape is described by its boundaries. Mathematically, it is convenient to study the deformation of domains by identifying each point of the original domain $\Omega$ with a point of the deformed domain $\Omega'$. This identification is required to be a diffeomorphism $h$. In order to assume no knowledge about $\Omega'$, one extends $h : \Omega \to \mathbb{R}^N$, such that $h$ is a diffeomorphism from $\Omega$ to $h(\Omega)$. One can construct a one parameter family $h(t, x)$ of such diffeomorphisms, parametrized by $t$, requiring that $h(0, x) : \Omega \to h(0, \Omega)$ acts like the identity of $\mathbb{R}^N$. The domain at a time $t$ is defined

to be $\Omega(t) = h(t, \Omega)$ and one clearly has $\Omega(0) = \Omega$. One can then differentiate various quantities related to the domains $\Omega(t)$ with respect to $t$ assuming that $h(t, x)$ is sufficiently differentiable when viewed as a map $h : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}^N$. For further details on this construction, we refer the reader to the excellent book by D. Henry [38].

By the method outlined above, it can be shown that simple eigenvalues of the Laplace-Beltrami operator for domains with Robin boundary conditions are differentiable at $t = 0$. This also holds for degenerate eigenvalues. Moreover, using the above techniques, it is possible to show that domains in $\mathbb{R}^N$ almost never have degenerate eigenvalues, except the zero eigenvalue for the Neumann problem, whose multiplicity is equal to the number of connected components of the domain. A similar caveat holds for Robin boundary conditions. This is shown by establishing that almost no diffeomorphism $h$ maps a domain into a domain with degenerate eigenvalues. Once again, we refer the reader to [38].

## 6.2 General Formulation

In this section, we propose a class of problems that we consider to be numerical realizations of infinitesimal inverse spectral geometry. We do not claim that this class of problems describes the totality of infinitesimal inverse spectral geometry, but merely that it is wide enough to be of interest. For this reason, the following discussion should not be seen as a strict mathematical definition, but rather as a list of guiding principles. Later sections provide examples of the application of those guidelines.

Since numerical methods can only handle finitely many numbers, it is impossible to numerically study the spectral geometry of shapes that possess infinitely many degrees of freedom. Similarly, one cannot consider the full spectrum of elliptic operators defined on such geometries as it is an infinitely long sequence. Consequently, our study is limited to the attempt to identify shapes given by a finite number of degrees of freedom from a finite number of eigenvalues of some elliptic operator defined on them. Let $n_{dof}$ and $n_{ev}$ denote the number of such degrees of freedom and the number of considered eigenvalues, respectively. In order to avoid unnecessary complications with constraints, we consider that the shape degrees of freedom take values in $\mathbb{R}^{n_{dof}}$. Strictly speaking, the studied part of the spectrum takes values in a potentially quite complicated subset of the nonnegative part of $\mathbb{R}^{n_{ev}}$. For the sake of simplicity, we shall ignore such technicalities and simply say that the spectrum takes values in $\mathbb{R}^{n_{ev}}$. Both $\mathbb{R}^{n_{dof}}$ and $\mathbb{R}^{n_{ev}}$ are equipped with the standard Euclidean metric, denoted $d(\cdot, \cdot)$. This choice is one of convenience, rather than of deep significance. It is quite possible that there are more natural notions of distance available.

74

The relationship between the set of degrees of freedom $\mathbb{R}^{n_{dof}}$ and the geometries it parametrizes is given by the *construction map* $C$. Let $\mathcal{G}$ denote the set of shapes obtained from $\mathbb{R}^{n_{dof}}$ via $C$, where the letter $\mathcal{G}$ stands for *geometries*. That is, $C : \mathbb{R}^{n_{dof}} \to \mathcal{G}$ is bijective by definition. In the present, the elements of $\mathcal{G}$ are supposed to be compact Riemannian manifolds. In more general settings, they could be taken to be of any other kind of object upon which an inverse spectral problem can be defined, such as potentials in the Schrödinger equation or spectral graphs. We insist that the elements of $\mathcal{G}$ must be sufficiently similar to one another for the procedure outlined here to make sense. For instance, if a particular element of $\mathcal{G}$ is a compact $N-$dimensional Riemannian manifold without boundary, all elements of $\mathcal{G}$ must also be $N-$dimensional Riemannian manifolds without boundary. We shall always identify points in $\mathbb{R}^{n_{dof}}$ with their image in $\mathcal{G}$. That is, we allow ourselves to say that $P \in \mathbb{R}^{n_{dof}}$ is a shape, when rigorously we should say that $C(P) \in \mathcal{G}$ is the shape corresponding to $P$.

**Remark 6.2.1.** *Distinct elements of $\mathcal{G}$ can be isometric. Let $[\mathcal{G}]$ be the set of equivalence classes of isometric elements of $\mathcal{G}$ and let $\pi : \mathcal{G} \to [\mathcal{G}]$ be the natural projection. The map $\pi \circ C$ needs not to be injective. We do not require that a given shape is encoded in a unique way by the degrees of freedom $\mathbb{R}^{n_{dof}}$. In other words, $\mathbb{R}^{n_{dof}}$ is not a coordinate system for $[\mathcal{G}]$. This at first seems to be a major problem in our proposal. It is not so. Recall that our motivation to study inverse spectral geometry is to obtain a method to specify shapes in a unique way. We are thus seeking a (possibly local) coordinate system for $[\mathcal{G}]$. Consequently, if a parametrization of $[\mathcal{G}]$ is already available, inverse spectral geometry becomes useless. For example, describing disks in $\mathbb{R}^2$ by their spectra rather than by their radii is impractical in most situations. In that sense, our inability to always uniquely describe $[\mathcal{G}]$ in terms of $\mathbb{R}^{n_{dof}}$ reflects the current limitations of differential geometry.*

One then selects a certain number of operators acting on some function spaces defined on elements of $\mathcal{G}$. For instance, one could pick the Laplacians acting on 0 and 1$-$forms of a Riemannian manifold. If boundary conditions are needed, they are of course selected. The operators chosen at this point are those whose spectra will be studied afterwards. We then define the *spectral map* $Sp : \mathcal{G} \to \mathbb{R}^{n_{ev}}$ which maps a given shape in $\mathcal{G}$ to $n_{ev}$ eigenvalues in its spectrum. The explicit computation of this map is when numerical methods come into play. Since the spectrum will generically contain more eigenvalues than $n_{ev}$, one must choose which eigenvalues to include. For instance, one could simply include the $n_{ev}$ smallest eigenvalues, counting multiplicity. Other choices are of course possible. It is also advisable to exclude eigenvalues that are a priori the same for all elements of $\mathcal{G}$, as they encode no information about the elements of $\mathcal{G}$. The zero eigenvalue of the Laplace-Beltrami operator on a compact Riemannian manifold without boundary is a good example of such

a case. If the spectra of multiple operators are studied, $\mathbb{R}^{n_{ev}}$ has to contain all of them. For example, if one considers the first 10 eigenvalues of the Laplacian on $0-$forms and the first 5 eigenvalues of the Laplacian on $1-$forms, it follows that $n_{ev} = 15$. In other words, $n_{ev}$ denotes the total number of considered spectral degrees of freedom. The $n^{th}$ entry in a spectrum is denoted $\lambda_n$.

For reasons that will be stated later, it is not always convenient to directly work with the spectrum of the operators. Instead, we introduce a *spectral adjustment function* $Adj : \mathbb{R}^{n_{ev}} \to \mathbb{R}^{n_{ev}}$. When its domain is restricted to spectra, rather than just sequences in $\mathbb{R}^{n_{ev}}$, this function is required to be invertible and smooth. We make extensive use of an adjustment function that sends nonzero eigenvalues to their multiplicative inverses, interchanging the spectrum of the Laplacian with that of the corresponding Green's operator. The $n^{th}$ entry in an adjusted spectrum will also be denoted $\lambda_n$. This slight abuse of notation is harmless, as context will dictate the meaning of $\lambda_n$. Similarly, we shall refer to the adjusted spectrum simply as the spectrum.

Finally, we define the *overall spectral map* $\sigma$ as the composition of the above. This is represented by the following diagram:

$$\mathbb{R}^{n_{dof}} \xrightarrow{\;\;C\;\;} \mathcal{G} \xrightarrow{\;\;Sp\;\;} \mathbb{R}^{n_{ev}} \xrightarrow{\;\;Adj\;\;} \mathbb{R}^{n_{ev}}$$
$$\sigma = Adj \circ Sp \circ C$$

This construction allows us to reduce the study of infinitesimal inverse spectral geometry to the study of $\sigma : \mathbb{R}^{n_{dof}} \to \mathbb{R}^{n_{ev}}$. We suppose this map to always be sufficiently differentiable for our purposes. The results on the continuity and differentiability of the eigenvalues reviewed in Section 6.1 guarantee that this hypothesis is reasonable. Given the above setting and terminology, the following definition sets the bounds of the setting for infinitesimal ISG that we consider. Once again, it is restricted when compared to the most general formulation possible and is to be taken as a guiding principle.

**Definition 6.2.1** (Infinitesimal ISG)**.** *Let $A, B \in \mathbb{R}^{n_{dof}}$. $A$ shall denote the starting shape, while $B$ shall denote the target shape. Suppose that $\sigma(A) \neq \sigma(B)$. The goal of infinitesimal inverse spectral geometry is to trace a continuous, parametrized path $P(t) \in \mathbb{R}^{n_{dof}}$ for $t \in [a, b]$, such that, for any pair $(A, B)$ :*

*(a) $P(a) = A$*

*(b)* $\sigma(P(b)) = \sigma(B)$

*(c)* $\sigma(P(t)) \neq \sigma(B)$ *for* $t \in (a, b)$

*(d)* $\frac{d}{dt}P(t)$ *is determined solely by* $\sigma(B)$ *and local data about* $\sigma$

*(e)* $d(\sigma(P(t')), \sigma(B)) \leq d(\sigma(P(t)), \sigma(B))$ *for* $t' > t$

*A systematic way of constructing such paths is called a spectrally full method of infinitesimal ISG. A systematic way of constructing paths that satisfy all of the above conditions except for (b) is called a spectrally partial method of infinitesimal ISG. Moreover, if for all* $A, B \in \mathbb{R}^{n_{dof}}$, $P(b)$ *and* $B$ *map into isometric shapes via the construction map* $C$, *it is said that the method is completely successful. If* $P(b)$ *and* $B$ *map into isometric shapes only for some* $A, B \in \mathbb{R}^{n_{dof}}$, *the method is partially successful. Clearly, a spectrally partial method cannot be completely successful. As spectrally partial partially successful methods are the most common, we will refer to them as methods of infinitesimal ISG.*

**Remark 6.2.2.** *Condition (d) of the above definition is what makes the approach infinitesimal. At a given point* $P(t)$, *infinitesimal ISG is only allowed to know where it is in the space of spectra, where it wants to go in the space of spectra and how the spectrum changes in a small neighbourhood of* $P(t)$. *Condition (e) ensures that* $P(t)$ *always tries to get closer to the target spectrum. This is particularly important for numerical implementations, as will become clear later. Continuity of the paths is a condition that will necessarily be violated for numerical implementations, as the problem needs to be discretized. In a broader perspective, this condition might need to be relaxed overall, allowing for non-continuous paths not on numerical, but on fundamental grounds. Finally, recall that* $d(\cdot, \cdot)$ *was chosen to be the Euclidean distance out of convenience. It need not be so in general.*

Essentially, inverse spectral geometry is reduced to an optimization problem that we try to solve by a method that tries to obtain global solutions by piecing together local ones. Its main difference with usual optimization problems is that the optimized function is not held fixed. Indeed, one is free to choose how many eigenvalues are being analyzed. Thus, if given a starting shape $A$ and a target spectrum $\sigma(B)$ one fails to find a path from $A$ to $B$, one can try again with more eigenvalues.

## 6.3   Explicit Implementation

The abstract principles of the previous section may be implemented in many ways. In this section we propose two natural ways of doing so and compare them. We then suggest

certain possible improvements of those techniques. Recall that we suppose the spectral map $\sigma : \mathbb{R}^{n_{dof}} \to \mathbb{R}^{n_{ev}}$ is sufficiently differentiable for all the manipulations we do to be well defined.

Let $d^2(\cdot, \cdot)$ denote the square of the standard distance on any Euclidean space. Using $d^2(\cdot, \cdot)$ rather than $d(\cdot, \cdot)$ simplifies some expressions. Minimizing $d^2(\cdot, \cdot)$ is equivalent to minimizing $d(\cdot, \cdot)$, so condition (e) of Definition 6.2.1 can be expressed with either. Let $\mathcal{J}(P)$ be the Jacobian matrix of $\sigma(P)$ at a point $P \in \mathbb{R}^{n_{dof}}$. Let $\mathcal{J}^T$ and $\mathcal{J}^+$ denote its transpose and its Moore-Penrose pseudoinverse, respectively. A brief introduction to the pseudoinverse can be found in Appendix A.

**Definition 6.3.1** (Spectral Direction)**.** *In the notation of Definition 6.2.1, at a point $P(t)$, the desired spectral direction is*

$$v_\sigma = \sigma(B) - \sigma(P(t)) \tag{6.3}$$

*The notation for $v_\sigma$ does not include $B$ and $P(t)$, as they will always be clear from the context.*

We propose the following two implementations of infinitesimal inverse spectral geometry.

**Definition 6.3.2** (Gradient Method)**.** *The parametrized path $P(t) \subset \mathbb{R}^N$ from $A$ to $B$ satisfies the following differential equation*

$$\frac{d}{dt}P(t) = -\mathbf{grad}\left(d^2(\sigma(P(t)), \sigma(B))\right) \tag{6.4}$$

**Definition 6.3.3** (Pseudoinverse Method)**.** *The parametrized path $P(t) \subset \mathbb{R}^N$ from $A$ to $B$ satisfies the following differential equation*

$$\frac{d}{dt}P(t) = \mathcal{J}^+(P(t))v_\sigma \tag{6.5}$$

In order to simplify the comparison of the two methods, the gradient method can be reformulated as follows.

**Proposition 6.3.1** (Gradient Method)**.** *The gradient method can be equivalently formulated as*

$$\frac{d}{dt}P(t) = 2\mathcal{J}^T(P(t))v_\sigma \tag{6.6}$$

*Proof.* This result depends on the fact that we chose the distance on $\mathbb{R}^{n_{ev}}$ to be Euclidean, as it implies that

$$d^2(\sigma(P(t)), \sigma(B)) = \|v_\sigma\|_2^2 \tag{6.7}$$

Consequently, the gradient approach reduces to

$$\frac{d}{dt}P(t) = -\mathbf{grad}\left(\|v_\sigma\|_2^2\right) \tag{6.8}$$

Let $\frac{\partial}{\partial x^j}$ denote the partial derivative with respect to the $j^{th}$ shape degree of freedom in $\mathbb{R}^{n_{dof}}$. Similarly, let $\sigma^i(P)$ denote the $i^{th}$ eigenvalue of $\sigma(P)$. We also write $v_\sigma^i = \sigma^i(B) - \sigma^i(P)$. The $j^{th}$ component of the gradient is then expressed as

$$
\begin{aligned}
\frac{\partial}{\partial x^j}\|v_\sigma\|_2^2 = \frac{\partial}{\partial x^j}\|\sigma(B) - \sigma(P)\|_2^2 &= \frac{\partial}{\partial x^j}\sum_{i=1}^{n_{ev}}(v_\sigma^i)^2 \\
&= -2\sum_{i=1}^{n_{ev}}(\sigma^i(B) - \sigma^i(P))\frac{\partial}{\partial x^j}\sigma^i(P) \\
&= -2\sum_{i=1}^{n_{ev}}v_\sigma^i \mathcal{J}_{ij}
\end{aligned}
\tag{6.9}
$$

Consequently, the gradient method can be recast as

$$\frac{d}{dt}P(t) = 2\mathcal{J}^T(P(t))v_\sigma \tag{6.10}$$

This completes the proof. $\qquad\square$

**Remark 6.3.1.** *Notice that since we are free to redefine the parameter $t$, the overall factor of 2 in Proposition 6.3.1 can be eliminated. In fact, any such positive factor can be eliminated or introduced as we see fit. This fact will be used later to speed up numerical computations.*

From the above remark, it can be seen that the difference between the two proposed methods is the difference between the transpose and the pseudoinverse of $\mathcal{J}$, up to positive scalar factors. It will sometimes be convenient to consider both methods simultaneously. To do that, let $\mathcal{J}^\bullet$ denote both $\mathcal{J}^T$ and $\mathcal{J}^+$. It is straightforward to show that both methods satisfy conditions $(a), (d)$ and $(e)$ of Definition 6.2.1. While condition $(b)$ will not be satisfied for all $A$ and $B$, it is to be expected that some pairs of initial and target shapes, in particular extremely close ones, can be connected by a path built by the proposed

79

methods. Both methods are thus expected to be at least spectrally partial and partially successful. Numerical experiments show that this is indeed the case.

Since both the gradient and pseudoinverse methods are not sure to succeed, it is of interest to explore the situations in which they stop moving. We consider that a method has stopped moving if one has $\frac{d}{dt}P(t) = 0$ for $v_\sigma \neq 0$. Clearly, this can only happen when $P(t)$ is not isospectral to $B$. In other words, a method gets stuck at $P_0 \in \mathbb{R}^{n_{dof}}$ if $v_s$ is in the kernel of $\mathcal{J}^\bullet(P_0)$. As explained in Appendix A, $ker(\mathcal{J}^T) = ker(\mathcal{J}^+)$. Thus, both the pseudoinverse and the gradient methods get stuck at exactly the same elements of $\mathbb{R}^{n_{dof}}$. This does not mean that the two methods have the same failure rate. Indeed, the fact that the two methods stop moving at the same points does not mean that they encounter those points with the same frequency. Put differently, it only implies that switching from one method to the other at a problematic point will not improve the situation. This is similar to a sighted and a blind man trying to find their way out of a maze. Neither the sighted nor the blind can walk through walls, yet the sighted hits walls less often.

The pseudoinverse method has an interesting advantage over the gradient one, as it can be used to define secondary objectives for the optimization process. As explained in Appendix A, the pseudoinverse can be used to define a projection operator onto $ker(\mathcal{J})$ via

$$P_{ker(\mathcal{J})} = \mathbb{1} - \mathcal{J}^+\mathcal{J} \tag{6.11}$$

Since vectors that lie in $ker(\mathcal{J})$ represent directions in $\mathbb{R}^{n_{dof}}$ that do not change the spectrum, one is free to move along such vectors, at least infinitesimally, without disturbing the spectrum. Thus, if one has a preferred direction $s_p$ to move in the space of shapes, it is possible to force the pseudoinverse approach to move in it as follows;

$$\frac{d}{dt}P(t) = \mathcal{J}^+(P(t))v_\sigma + P_{ker(\mathcal{J})}s_p \tag{6.12}$$

This idea is exploited in the field of inverse kinematics, which studies the positioning of multi-jointed robotic arms. A review of this field can be found in [12]. More specifically, inverse kinematics studies how to move the end-effector of the arm, the part that accomplishes given tasks, from an initial to a desired position. One of the possible approaches to solve this problem is entirely analogous to the pseudoinverse approach that we propose. Since the engineers that design robotic arms are fully aware of the possible configurations of the arm, they know which positions are pathological and are to be avoided. They can thus steer the system away from such configurations using secondary objectives $s_p$ for the pseudoinverse method.

In our case, there is no obvious choice of a preferred direction as we are ignorant of the exact properties of the spectral map $\sigma$. We instead rely on the linear algebraic structure of the pseudoinverse method to attempt to define a secondary objective that would promote the avoidance of problematic points. Consider a point $P \in \mathbb{R}^{n_{dof}}$ at which $v_\sigma \in ker(\mathcal{J}^+(P))$. It is expected that this pathological behaviour would not be exactly valid for points in a neighbourhood of $P$. Instead, it is likely that the pathological behaviour arises in a continuous fashion. However, the dimension of the kernel is a discrete number and thus cannot change continuously. It is thus necessary to choose a way to measure the rank of $\mathcal{J}$ in a continuous way. This is made possible by the notion of effective rank. Let $n_{min} = \min(n_{ev}, n_{dof})$ and let $\{s_i\}_{i=1}^{n_{min}}$ denote the multi-set of singular values of $\mathcal{J}$. Since singular values are by construction positive, one can introduce a normalized version of this multi-set $\{\hat{s}_i\}_{i=1}^{n_{min}}$ so that $\sum_{i=1}^{n_{min}} \hat{s}_i = 1$. This makes $\{\hat{s}_i\}_{i=1}^{n_{min}}$ look like a probability distribution. One can define an entropy $H(\mathcal{J})$ over this multi-set:

$$H(\mathcal{J}) := - \sum_{i=1}^{n_{min}} \hat{s}_i \log(\hat{s}_i) \tag{6.13}$$

In the above expression, it is taken that $0 \log(0) = 1$. The effective rank of $\mathcal{J}$ is then defined to be

$$R_e(\mathcal{J}) = \exp(H(\mathcal{J})) \tag{6.14}$$

For mathematical motivations of this definition, see [19, 65]. Physical motivations can be found in [20]. Consider the case of maximal entropy: all of the $\{\hat{s}_i\}_{i=1}^{n_{min}}$ are equal and nonzero. In that case, the rank of $\mathcal{J}$ is equal to $n_{min}$, as expected. Similarly, when $p$ of the $\hat{s}_i$ are nonzero and equal and the remaining ones are zero, the effective rank is $p$. In the extreme case $p = 1$, only one $\hat{s}_i$ is nonzero and must be equal to one. In [65], it is shown that

$$1 \leq R_e(\mathcal{J}) \leq \mathrm{rank}(\mathcal{J}) \leq n_{min} \tag{6.15}$$

Moreover, the first two inequalities may only hold with equality in the special cases described above. For our purposes, the maximization of the effective rank may serve as a secondary objective via

$$s_p = \mathbf{grad}(R_e(\mathcal{J})) \tag{6.16}$$

81

The purpose of this secondary objective is to try to avoid points in $\mathbb{R}^{n_{dof}}$ at which $\mathcal{J}$ has a rank lower than usual as much as possible. Since singular values of $\mathcal{J}$ are expected to vary continuously with $P \in \mathbb{R}^{n_{dof}}$, a drop in rank should be predictable *before it happens* from a drop in effective rank. As we expect pathological points where $v_\sigma \in ker(\mathcal{J})$ to be associated with a lower than usual rank, maximizing the effective rank as a secondary objective could be a strategy for avoiding them. It is not clear at this point if using this secondary objective produces better results in numerical implementations. It does however produce different results while using significantly more computational resources. Further investigation is required.

## 6.4   Numerical Implementation

The gradient and pseudoinverse methods can be implemented numerically. Assume that the spectral map $\sigma(P)$ can be computed pointwise. In general, this entails the usage of some numerical approximation of the problem. For example, if the objects in $\mathcal{G}$ are manifolds, finite element methods may be used. For now, the precise nature of such methods is of no importance and it is sufficient to assume that they are available. The Jacobian matrix $\mathcal{J}$ is obtained by a finite difference approximation of the gradient. It is expected that a numerical approximation of $\sigma$ would have some numerical error due to the approximations used. Even if this error is small in magnitude, it may vary much faster than the eigenvalues, significantly affecting the Jacobian. For this reason, it is advantageous to compute the gradient multiple times using various values of the step size and take the average result.

It is of course necessary to discretize the steps taken in the space of degrees of freedom $\mathbb{R}^{n_{dof}}$. The gradient and pseudoinverse methods will not be implemented in the form given by Definition 6.3.3 and Proposition 6.3.1. Instead, it is useful to be able to control the size of the step taken, so that the actual magnitude of the gradient does not influence the speed of the algorithm. This point will become clear soon. Let $\mathbf{N}(v)$ denote a normalized version of the vector $v$ and set $\mathbf{N}(0) = 0$. Let $S$ denote a positive step size. It will change as the numerical implementation runs. The time parameter $t$ is discretized to number the steps taken by the algorithm. Both methods will be expressed as

$$P(t+1) = P(t) + S\mathbf{N}\left(\mathcal{J}^\bullet(P(t))v_\sigma\right) \tag{6.17}$$

The above discretization can in principle be implemented numerically. However, as it is stated now, the method has multiple easy to perceive problems. Choose some $A, B \in \mathbb{R}^{n_{dof}}$ and, for now, assume that the method studied can trace a continuous path between $A$ and

$B$, in the sense of Definition 6.2.1. It is clear that, for a given $S$, it is unlikely that the iteration of Equation (6.17) would bring $P(t)$ exactly from $A$ to $B$. Indeed, it is not even guaranteed that the discretization brings one within a given ball of a given small radius centred at $B$. In a more general setting, it is not guaranteed that there is a $t_1$ such that the spectra $\sigma(P(t_1))$ and $\sigma(B)$ match. Of course, if $S$ is taken small enough, one can expect this not to be a problem. However, small step size will increase the time necessary to compute the path $P(t)$. One thus must simultaneously introduce two things: a tolerance threshold at which it is considered that two spectra are equal and a mechanism that controls the step size $S$ so that steps sufficiently small to get $\sigma(P(t))$ within tolerance of $\sigma(B)$ can be taken. The tolerance threshold shall be denoted $\varepsilon_\sigma$. It is considered that the algorithm has converged if, for some $t_1$, one has

$$d(\sigma(P(t_1)), \sigma(B)) \leq \varepsilon_\sigma \tag{6.18}$$

One can then stop the iteration, effectively setting $b = t_1$. The steps size $S$, and the whole process, is controlled by the following algorithm.

**Definition 6.4.1** (Race Car Algorithm). *Set some initial step size $S = S_0$ and two positive factors $u > 1$ and $d < 1$, where $u$ stands for "up" and $d$ for "down". Set $S_{min}$ to be the minimal allowed step size. Those values depend upon the problem considered and are determined by trial and error. $S$ will vary during the calculation, but $u$ and $d$ will not. Let $P(t)$ be the current shape and let $P(t+1)$ be the next shape to be determined. The discretized path between $A$ and, hopefully, $B$ is constructed by going through the following steps.*

1. *If $d(\sigma(P(t)), \sigma(B)) \leq \varepsilon_\sigma$ or $S < S_{min}$, stop the iterations.*

2. *Compute $\mathcal{J}(P(t))$.*

3. *Compute $P(t+1)$ using Equation (6.17).*

4. *If $d(\sigma(P(t+1)), \sigma(B)) \leq d(\sigma(P(t)), \sigma(B))$, $P(t+1)$ is accepted. Increase the step size by setting $S \to Su$. Go back to step 1, replacing all instances of $t$ with $t+1$.*

5. *Otherwise, one has $d(\sigma(P(t+1)), \sigma(B)) > d(\sigma(P(t)), \sigma(B))$ and the algorithm moved further away from the goal. Set $P(t+1) = P(t)$, rejecting the step. Decrease the step size by setting $S \to Sd$. Go back to step 1, replacing all instances of $t$ with $t+1$.*

The name Race Car algorithm was coined by Achim Kempf [46]. It refers to the fact that the algorithm is always either accelerating or decelerating exponentially fast, as if driven by a crazy race car pilot who always either has his foot on the accelerator or on the brakes. The algorithm satisfies both the need for a speedup when the current and target spectra are far away from one another and the need for a more careful manoeuvring when the algorithm has almost converged to an isospectral geometry. In all of the following, we choose $u = 1.1$ and $d = 0.7$. Notice that the acceptance criterion for a step is the discretized version of property $(e)$ of Definition 6.2.1.

The Race Car algorithm can stop for two reasons. The first is the the achievement of the desired result: a shape isospectral to $B$. The second is the step size being reduced below acceptable levels. This indicates that the algorithm is stuck. A stuck algorithm needs not to have failed completely. Consider the following extreme case. Due to numerical error (in which we include approximation error), if one sets the tolerance threshold $\varepsilon_\sigma$ too low, no $P(t)$ will ever trigger the condition $d(\sigma(P(t)), \sigma(B)) \leq \varepsilon_\sigma$. Thus, the algorithm will always get stuck, even if $P(t)$ arrives very close to $B$.

Of course, the algorithm may also get stuck quite far from $B$. In order to properly discuss the success and failure rates of the algorithm, one must define a notion of distance on the set of isometry equivalence classes $[\mathcal{G}]$. In other words, a notion of distance on the set of shapes $\mathcal{G}$ such that isometric shapes are a distance zero apart is needed. The second formulation is easier to implement, as the construction map $C$ gives one access to the shapes, but not to the equivalence classes of isometric shapes. Identifying shapes with their coordinates on $\mathbb{R}^{n_{dof}}$, let $d_\mathcal{G} : \mathbb{R}^{n_{dof}} \times \mathbb{R}^{n_{dof}} \to \mathbb{R}$ denote such a map. Using this distance, define a tolerance threshold $\varepsilon_\mathcal{G}$. Once the Race Car algorithm has stopped one has to verify if the final shape $P(b)$ is sufficiently close to $B$. We consider that $P(b)$ is isometric to $B$ if

$$d_\mathcal{G}(P(b), B) \leq \varepsilon_\mathcal{G} \tag{6.19}$$

In that case, the calculation is a success, as a path from $A$ to $B$, or an isometric $P(b)$ was traced. Otherwise, it is a failure, either because $P(b)$ is not isospectral to $B$ and the algorithm is stuck or $P(b)$ is isospectral but not isometric to $B$. The two different ways to fail have to be considered independently when analyzing success and failure rates. The second type of failure indicates the presence of isospectral non-isometric shapes in $\mathcal{G}$. This type of failure is in principle understood. All other sources of failure are of first type and their precise nature is yet unexplained. Their presence points to either serious numerical errors or fundamental obstructions to infinitesimal inverse spectral geometry. This type of failure should thus be carefully explored, both for the sake of eliminating

84

potential numerical difficulties and, more importantly, to understand the phenomenology of fundamental obstructions to infinitesimal inverse spectral geometry. Such an exploration would take a great deal of computational time and is outside the scope of this thesis. Still, it is an important thing to consider in further work, as knowing the phenomenology of failures could indicate how to overcome them.

The tolerance thresholds $\varepsilon_\sigma$ and $\varepsilon_\mathcal{G}$ must be mutually consistent. If we consider that $P(b)$ and $B$ are effectively isometric when $d_\mathcal{G}(P(b), B) \leq \varepsilon_\mathcal{G}$, they must also be effectively isospectral $d(\sigma(P(b)), \sigma(B)) \leq \varepsilon_\sigma$. In practice, it is much easier to impose a criterion on differences between shapes than on differences between spectra. There are two distinct reasons for this. First, shapes are easier to visualize than spectra and it is thus easier to decide which shapes are considered equivalent. Second, the number of degrees of freedom of shapes is fixed, while the number of eigenvalues may be altered as needed, potentially requiring a redefinition of $\varepsilon_\sigma$. To an extent, this suggests to try to do without $\varepsilon_\sigma$, relying solely on the minimal step size $S_{min}$ to stop the Race Car algorithm. The advantage of this approach is that it allows $P(b)$ to get as close to $B$ as the algorithm allows it. On the other hand, it also implies an increase in computation times, as the algorithm is required to compute extra iterations until the step size decreases below $S_{min}$. In our numerical experiments, we use an intermediary approach, where $\varepsilon_\mathcal{G}$ is fixed to a reasonable value while $\varepsilon_\sigma$ is chosen to be too low, so that many effectively isometric shapes are treated to be non-isospectral. Then, when a sufficient amount of data on the studied problem is collected, $\varepsilon_\sigma$ is increased to the minimum value that ensures that all effectively isometric shapes are also effectively isospectral.

## 6.5   Which Method to Use?

Up to this point, we discussed the gradient and pseudoinverse methods as equally valid. This ends here, as the goal of this section is to convince the reader of the superiority of the pseudoinverse approach. This is not immediately obvious as, given the Jacobian, it is of course faster to compute its transpose than its pseudoinverse. Thus, one could expect the gradient method to converge faster. This is not so. The power of the pseudoinverse method lies in the fact that it tries to take the most direct path in $\mathbb{R}^{n_{ev}}$ from $\sigma(A)$ to $\sigma(B)$. Computationally, at the very least in the regimes we explored, this is of tremendous benefit. While every single step takes longer to compute, the overall number of steps necessary to get from $A$ to a shape $P(b)$ isospectral to $B$ is much lower.

The above situation can be illustrated by a simple example. Let $\mathbb{R}^{n_{dof}} = \mathbb{R}^2 = \mathbb{R}^{n_{ev}}$. Consider the following toy model of a spectral map:

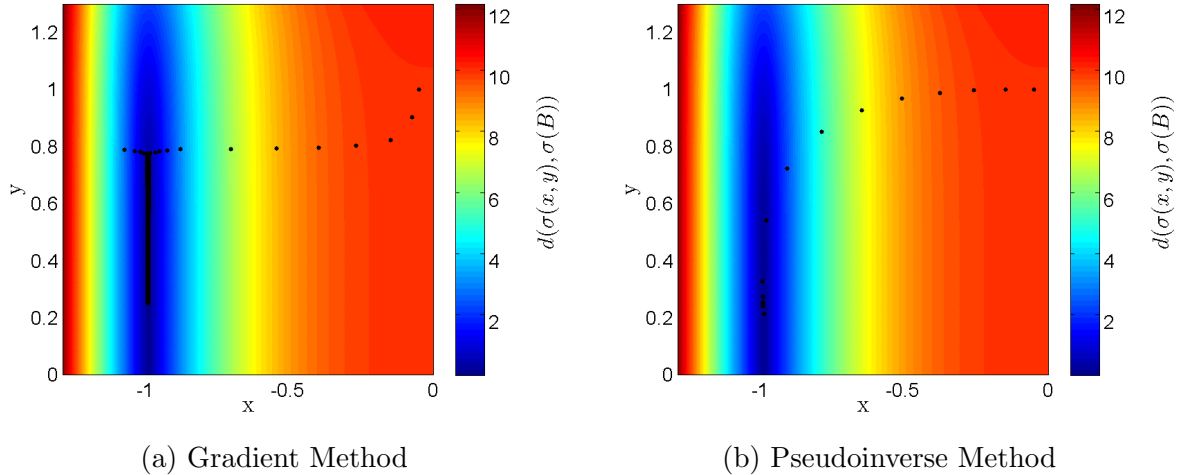(a) Gradient Method    (b) Pseudoinverse Method

Figure 6.1: Superiority of the pseudoinverse method. Paths taken by the gradient and pseudoinverse methods for the toy model defined by Equation (6.20) with $A = (-0.05, 1)$ and $B = (-1, 0.25)$. The colour gradients represent $d(\sigma(x, y), \sigma(B))$. The gradient approach needs 48957 steps and approximately 6 minutes of computation to get $\sigma(P)$ within $\varepsilon_\sigma = 10^{-4}$ of $\sigma(B)$. On the same computer, the pseudoinverse method accomplishes this in 26 steps and 0.24 seconds.

$$\sigma(x, y) = (10x^3, y^3) \tag{6.20}$$

Clearly, $\sigma$ is invertible on all of $\mathbb{R}^{n_{dof}}$. Set $A = (-0.05, 1)$ and $B = (-1, 0.25)$. One can try to use the gradient and pseudoinverse methods to trace paths from $A$ to $B$ using the Race Car algorithm. Figure 6.1 illustrates the difference in performance between the gradient and pseudoinverse methods.

The reason for the dramatic difference in performance is the fact that the gradient method follows the steepest descent and is thrown into a valley with steep walls centred about $x = -1$. At the bottom of the valley, the negative gradient of $d(\sigma(P), \sigma(B))$ points towards $B$. However, the discretized gradient method never quite hits $x = -1$. Thus, $P(t)$ is likely to lie on the steep wall of the valley. At such points, the gradient is mostly directed towards the centre of the valley, with only a small component in the direction of $B$. Correspondingly, $P(t + 1)$ will be on the opposite wall of the same valley, where a similar situation arises. So, instead of going in the direction of $B$, the gradient approach concentrates its efforts on jumping from one wall to another, a rather unproductive activity,

both in optimization and in real life. Put in the same situation, the pseudoinverse chooses the appropriate direction, regardless of the high slope of the valley walls. As promised, the pseudoinverse method reaches $B$ in a significantly lesser number of steps. The meandering behaviour of the gradient method is a well known problem in optimization literature and solutions to it, such as adding friction terms, have been proposed [63]. Instead of implementing such methods, we shall simply limit our investigation of numerical infinitesimal inverse spectral geometry to the pseudoinverse method.

## 6.6 Star-Shaped Domains in $\mathbb{R}^2$

In this section, we apply infinitesimal inverse spectral geometry in the sense of Definition 6.2.1 to a particular class of star-shaped domains in $\mathbb{R}^2$ defined by exponentials of Fourier Series. The spectrum of interest is that of the Laplace-Beltrami operator with Dirichlet boundary conditions. First, we establish the class of shapes that we study and exhibit certain of their properties.

### 6.6.1 Partial Exponential Fourier Domains

Let $(r, \phi)$ be the standard polar coordinates in the standard $\mathbb{R}^2$ plane equipped with the standard metric. We are going to study domains star-shaped about the origin, whose boundary is expressed as a radius function $R(\phi)$. In order to make this function dependent upon a finite number of degrees of freedom, we shall choose $R(\phi)$ as an exponential of a truncated Fourier series:

$$R(\phi) = 0.1 + 0.9 \exp\left(C_0 + \sum_{n=1}^{M} [C_i \cos(n\phi) + S_i \sin(n\phi)]\right) \tag{6.21}$$

The choice of the positive constants 0.1 and 0.9 is of course arbitrary. We chose those parameters so that a disk of radius 1 is obtained by setting all the coefficients $C_i$ and $S_i$ to zero. The exponential function ensures that $R(\phi)$ is positive. The additive constant 0.1 is necessary for numerical implementations, as otherwise the discretization of the domain fails due to non-adjacent parts of the boundary being too close to each other. The coefficients $\{C_i\}_{i=0}^{M}$ and $\{S_i\}_{i=1}^{M}$ are taken to be the degrees of freedom of the shape. In the sense of Definition 6.2.1, they are the space $\mathbb{R}^{n_{dof}}$, and $R(\phi)$ is the construction map associated to some $P \in \mathbb{R}^{n_{dof}}$. We will indicate this dependence by a subscript: $R_P(\phi)$. Consequently, $\mathcal{G}$

is the set of all domains generated by Equation (6.21). In the particular cases we explore, some coefficients of order lower than $M$ will be fixed to be non-dynamical. Except for the coefficient $C_0$, the degrees of freedom will always come in sine-cosine pairs of the same frequency. This allows for rotations of the studied shapes, as is best explained by the following well known trigonometric identity.

**Lemma 6.6.1** (Well-Known Trigonometric Identity). *For any $a, b \in \mathbb{R}$, there exists $\delta \in [0, 2\pi)$ such that*

$$a\cos(x) + b\sin(x) = \sqrt{a^2 + b^2}\sin(x + \delta) \tag{6.22}$$

**Remark 6.6.1.** *Trigonometric identities are not normally stated with such pomp. However, we give special treatment to the above result as it is used quite a few times in the following discussion. We shall not need explicit expressions for $\delta$.*

By virtue of Lemma 6.6.1 including sines and cosines in pairs of same frequency in the sum ensures that Equation (6.21) allows for the definition of all rotations of a given shape about the origin. One can explicitly obtain the rotated version of a domain by changing the coefficients via the following rule:

**Lemma 6.6.2** (Rotation of Shape). *A rotation $R_P(\phi - \delta)$ of $R_P(\phi)$ can be obtained via Equation (6.21) by modifying the coefficients as follows:*

$$
\begin{aligned}
C_0 &\to C_0 \\
C_n &\to C_n\cos(n\delta) - S_n\sin(n\delta) \\
S_n &\to C_n\sin(n\delta) + S_n\cos(n\delta)
\end{aligned}
\tag{6.23}
$$

*Proof.* The proof is trivial. $\qquad\square$

Similarly, one can obtain the rule for the reflexion of the shape with respect to the $x$ axis.

**Lemma 6.6.3** (Reflexion of Shape). *A reflexion of $R_P(\phi)$ with respect to the $x$ axis can be obtained via Equation (6.21) by modifying the coefficients as follows:*

$$
\begin{aligned}
C_0 &\to C_0 \\
C_n &\to C_n \\
S_n &\to -S_n
\end{aligned}
\tag{6.24}
$$

*Proof.* The proof is trivial. $\qquad\square$

The above two lemmata will be used to define a notion of distance $d_{\mathcal{G}}(\cdot, \cdot)$ between the shapes. We want for shapes equivalent by rotation and reflexion followed by rotation to be a distance zero apart, as they are isometric. We do not consider translation isometries. This is a minor problem whose solution we defer for now. Moreover, we want the distances given by $d_{\mathcal{G}}(\cdot, \cdot)$ to be easy to interpret, as this will facilitate the choice of the tolerance threshold $\varepsilon_{\mathcal{G}}$. Let $P \in \mathbb{R}^{n_{dof}}$ and let $[P] \subset \mathbb{R}^{n_{dof}}$ be the set of shapes equivalent to $P$ through rotations and reflexions followed by rotations, in the sense of lemmata 6.6.2 and 6.6.3. The distance between two shapes $P, Q \in \mathbb{R}^{n_{dof}}$ is defined as follows:

$$d_{\mathcal{G}}(P, Q) = \inf_{[P]} \left( \sup_{\phi} |R_P(\phi) - R_Q(\phi)| \right) \tag{6.25}$$

In other words, $d_{\mathcal{G}}(\cdot, \cdot)$ is the supremum norm distance between $R_P(\phi)$ and $R_Q(\phi)$, minimized over all shapes considered isometric to $P$. In numerical computations, one must discretize $\phi$ to determine the optimal angle $\delta$ by which one must rotate $R_P(\phi + \delta)$, or its reflexion. In our computations, $2\pi$ is divided in 1440 segments of equal length. The shapes we study usually have a diameter somewhere between 1 and 3. Thus, we consider that $\varepsilon_{\mathcal{G}} = 0.01$ is a reasonable choice for the shape equivalence threshold. Numerical experiments show that $\varepsilon_{\sigma} = 10^{-9}$ is a good a priori threshold for isospectrality when the spectrum of the Green's operator is used.

As mentioned earlier, the isometry group under which we require $d_{\mathcal{G}}(\cdot, \cdot)$ to be invariant is not the full isometry group of $\mathbb{R}^2$, as translations are missing from it. In most cases, this is a minor problem, as we only allow for a finite number of Fourier components in the description of the shape. Generically, to describe a translated version of the same shape, one would need more Fourier coefficients than those that we provide. As an extreme example, consider that only $C_0$ is allowed to be nonzero. Clearly, it is then impossible to describe anything but disks centred at the origin. Correspondingly, the description of a disk of radius $r$ centred at $\alpha \leq r$ will require more Fourier components than just the constant one. This state of affairs breaks down in one important case. Consider a disk of radius 1 centred at $(\alpha, 0)$ for $0 < \alpha \ll 1$. From Figure 6.2 and the law of sines, it is straightforward to show that in terms of $\phi$, the boundary of the disk is given by

$$R(\phi) = \frac{\sin(\phi + \arcsin(\alpha \sin \phi))}{\sin \phi} \approx 1 + \alpha \cos \phi \tag{6.26}$$

where the approximation is given by the first two terms of a Taylor series around $\alpha = 0$. Compare this to the following approximation of a valid boundary shape $\tilde{R}(\phi)$ for $C_1$ near zero
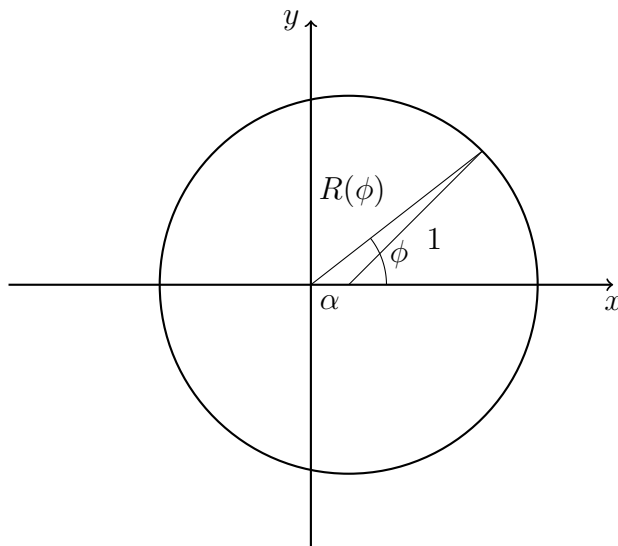
Figure 6.2: Translation of the unit disk centered at $(0,0)$ by $\alpha$ along the $x$ axis.

$$\tilde{R}(\phi) = 0.1 + 0.9 \exp(C_1 \cos \phi) \approx 1 + 0.9 C_1 \cos \phi \qquad (6.27)$$

Clearly, it is possible to choose $C_1$ so that $\tilde{R}(\phi)$ is a good approximation of a translated disk of radius 1. Effectively, this introduces translational isometries for which our method does not account. By Lemma 6.6.1, this discussion is valid for a disk of radius 1 centred at any point in $x \in \mathbb{R}^2$ such that $|x| \leq \alpha$. Consequently, the Fourier components of index 1 in Equation (6.21) will be considered pathological and their coefficients $C_1$ and $S_1$ will be fixed to be zero, to avoid the introduction of unaccounted for isometries.

## 6.6.2   Numerical Results

In order to conduct numerical experiments on the infinitesimal inverse spectral geometry of the partial exponential Fourier domains defined above, we must be able to compute the first $n_{ev}$ eigenvalues of the spectrum of the Laplace-Beltrami operator with Dirichlet boundary conditions. We do so by using finite element methods. A minimal introduction to finite elements can be found in Appendix B. As finite element methods may be quite subtle to implement from scratch, we use the freely available finite element solver FreeFem++. In the grand scheme of things, the results that we present here are preliminary and are to

be taken as a proof of concept more than as reliable data. We present success rates for a numerical implementation of the pseudoinverse method applied to a particular class of domains. We study how success rates change as functions of the number of eigenvalues considered by the algorithm and the initial distance $d_{\mathcal{G}}(A, B)$.

For the remainder of the discussion, we use the pseudoinverse method without secondary objectives and a spectral adjustment function that sends the nonzero eigenvalues of $\Delta$ to their multiplicative inverses and leaves zero eigenvalues unmodified. Thus, rather than studying the first $n_{ev}$ eigenvalues of the Laplace-Beltrami operator, we study the first $n_{ev}$ eigenvalues of its Green's operator. The pseudoinverse method is chosen as it is not prone to the meandering behaviour of the gradient method. No secondary objective was used as preliminary results on their usefulness were inconclusive. The spectral adjustment function was chosen because our preliminary results seemed to indicate that using the inverse eigenvalues is significantly more successful. More precisely, it seems that using the spectrum of the Green's operator reduces the chance of the Race Car algorithm to get stuck. Plausibly, this improvement is due to the fact that using multiplicative inverses of the eigenvalues of the Laplacian rather than the eigenvalues themselves modifies the relative importance of each eigenvalue. For the Laplacian, the low order eigenvalues have small magnitude and high order eigenvalues have large magnitude. This relationship is inverted for the spectrum of the corresponding Green's operator. Consequently, the algorithm favours changes of shape that bring low order eigenvalues closer to the target ones. Since low order eigenvalues correspond to long wavelength eigenfunctions, they are expected to only be sensitive to large changes in shape. Thus, prioritizing low order eigenvalues by using the spectrum of the Green's operator rather than that of the Laplacian encourages the algorithm to first fix the rough shape of the manifold and then worry about the smaller details. A more extensive investigation of this phenomenon would be of interest. In particular, it is possible that more refined spectral adjustment functions would yield better results.

The general setup for our numerical experiments is as follows. We fix the number of shape degrees of freedom $n_{dof}$ and choose to which Fourier coefficient in Equation (6.21) each axis of $\mathbb{R}^{n_{dof}}$ corresponds. Those will be called the dynamical Fourier coefficients. The other Fourier coefficients are called non-dynamical and are the same for all the shapes considered. We choose to fix the coefficients of orders $0, 1$ and $2$ to zero and let coefficients of orders $3$ through $5$ vary. In total, six shape degrees of freedom are chosen. Due to the trigonometric identity of Lemma 6.6.1, one of the coefficients is redundant, as it simply describes the orientation of the shape. It will thus be eliminated when the algorithm checks for isometry. Effectively, only five degrees of freedom remain.

Each time the algorithm runs, an initial shape $A$ and a target shape $B$ are randomly

generated. Let $r(-1,1)$ denote a function that generates uniformly distributed random numbers in $(-1,1)$. The dynamical Fourier coefficients of $A$ and $B$ are generated according to the following rule

$$
\begin{aligned}
C_i(B) &= M_B r(-1,1) \\
S_i(B) &= M_B r(-1,1) \\
C_i(A) &= M_A r(-1,1) + C_i(B) \\
S_i(A) &= M_A r(-1,1) + S_i(B)
\end{aligned}
\tag{6.28}
$$

where $M_B$ and $M_A$ are positive numbers that dictate the size of the coefficients. The magnitude of $M_B$ dictates how wide is the class of considered shapes. $M_A$, on the other hand, controls how close are the starting points to the target point.

Once the shapes are generated, the spectrum $\sigma(B)$ is computed and the Race Car algorithm implementation of the pseudoinverse method is used to attempt to trace a path from $A$ to $B$.

The first result we wish to state is a qualitative one. The pseudoinverse approach to inverse spectral geometry sometimes succeeds in finding the target shape, or a shape isometric to it. Success depends on the number of eigenvalues considered. An example of this is illustrated on Figures 6.3 and 6.4. Both figures present the shapes obtained by the pseudoinverse approach for the same pair of initial and target points $A$ and $B$. In Figure 6.3, $n_{ev} = 6$ is used and the algorithm fails to find the target shape. In Figure 6.3, on the other hand, $n_{ev} = 10$ is used and the algorithm succeeds.

The above observation leads to the following question. Given a fixed number of degrees of freedom, how many eigenvalues are needed to reliably find the target shape? From a simple matching of degrees of freedom argument, it is clear that one must have $n_{ev} \geq n_{dof}$ to even begin to hope that a bijection exists. Anything else is up to numerical experimentation. We have carried out such a numerical experiment for a class of shapes whose dynamical Fourier coefficients are those of orders 3 through 5, for a total of $6 - 1$ degrees of freedom, as explained above. The experiment consisted of the study of 3076 randomly generated pairs of initial and final shapes $A$ and $B$. The coefficients of $B$ were set using $M_B = 0.25$. Different values of $M_A$ were used to obtain different starting distances. For each pair $(A, B)$, the Race Car algorithm was ran 40 times, with different numbers of eigenvalues considered: $n_{ev} = 1...40$. The final shapes $P(b)$ obtained by the algorithm were compared to the target shapes $B$ an if $d_\mathcal{G}(P(b), B) \leq 0.01$, the run was declared a success. Otherwise, it was deemed a failure. Figure 6.5 presents the success rates as a function of $n_{ev}$ and of the starting shape distance $d_\mathcal{G}(A, B)$.
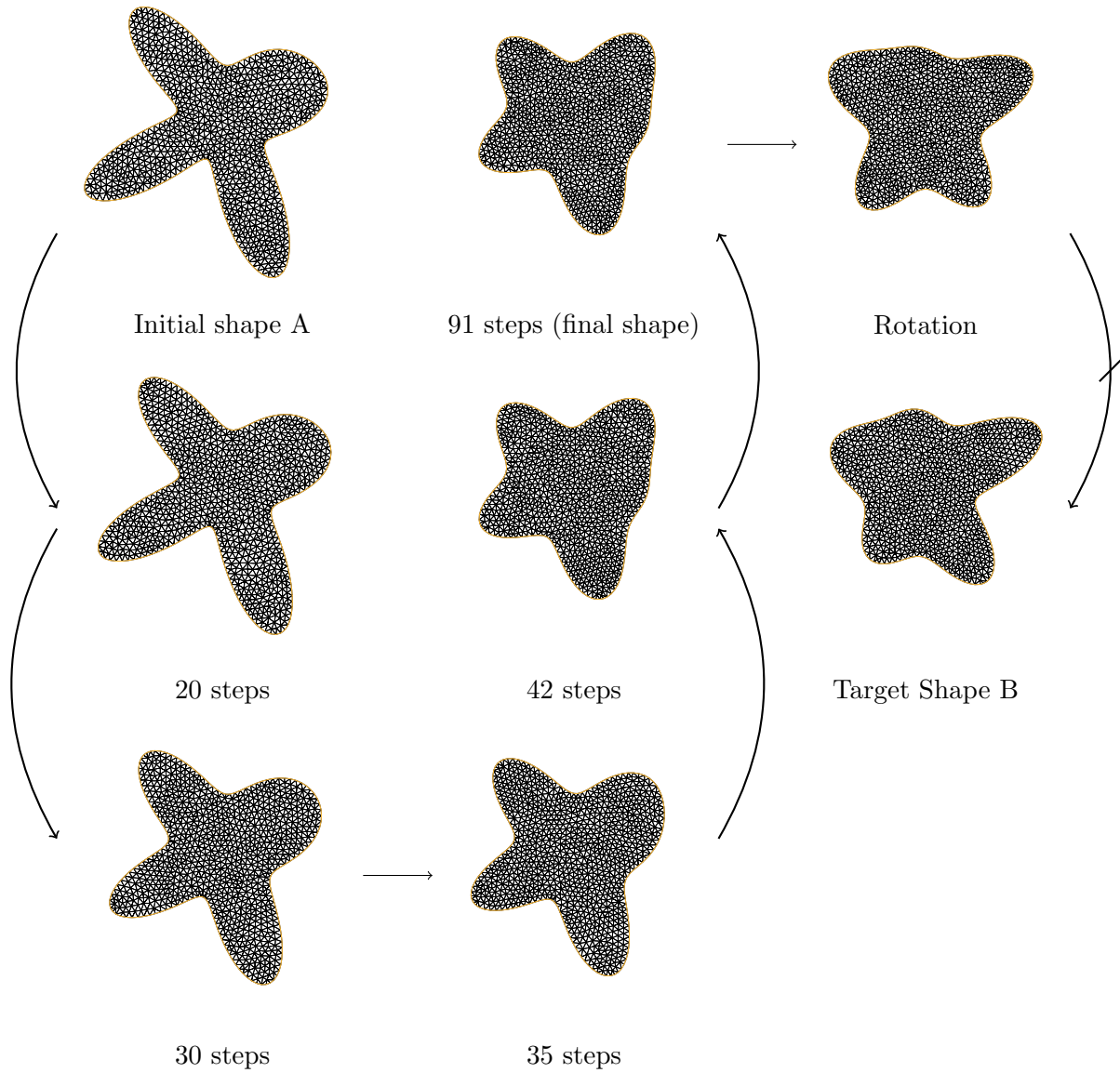
Initial shape A  91 steps (final shape)  Rotation

20 steps  42 steps  Target Shape B

30 steps  35 steps

Figure 6.3: Failure of the Race Car Algorithm for $n_{dof} = 6, n_{ev} = 6$. The final shape distance is $d_{\mathcal{G}}(P(93), B) = 0.176 > 0.01$. The final spectral distance is $d(\sigma(P(93)), \sigma(B)) = 4.8 \cdot 10^{-6} > 10^{-9}$.
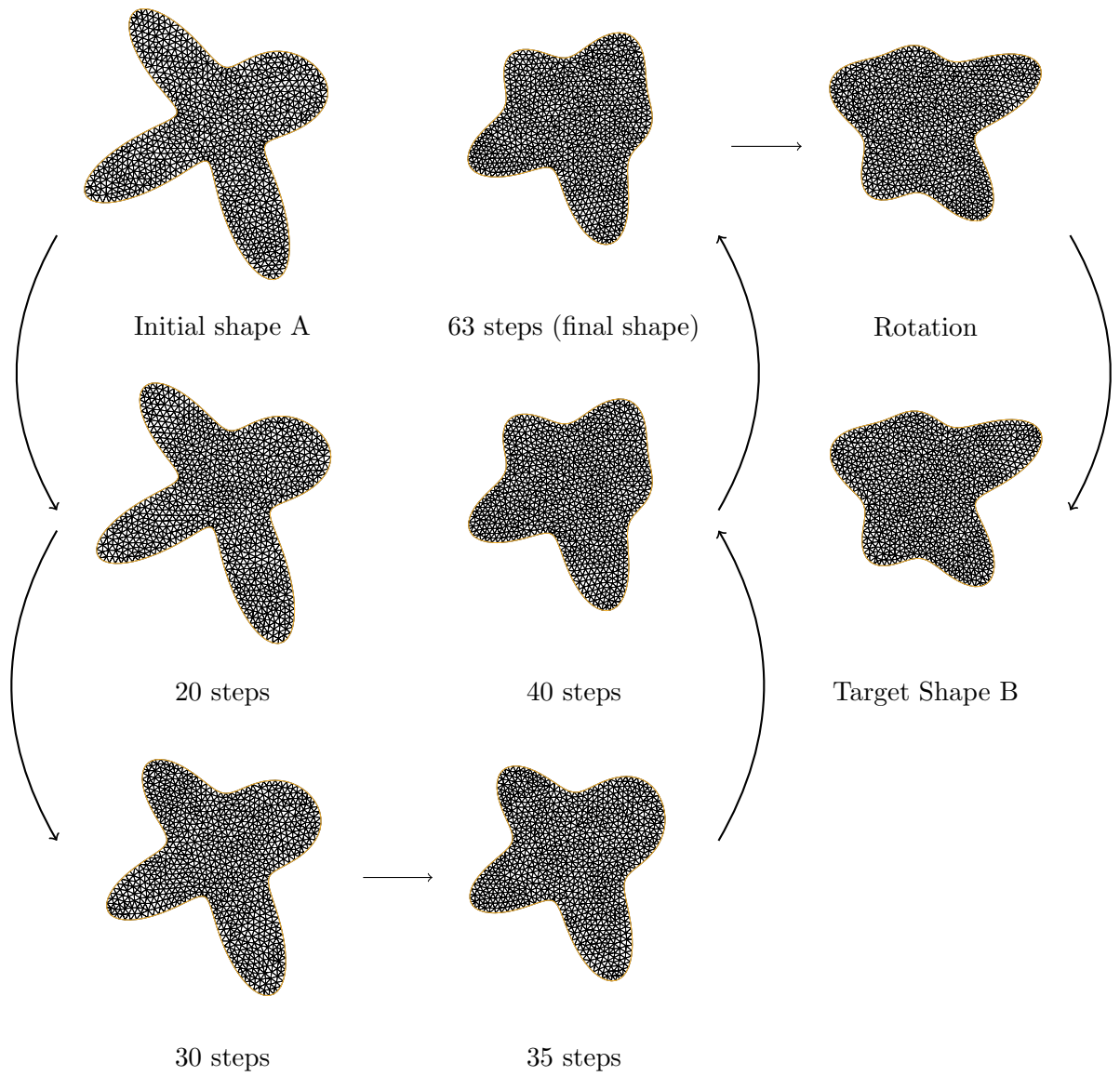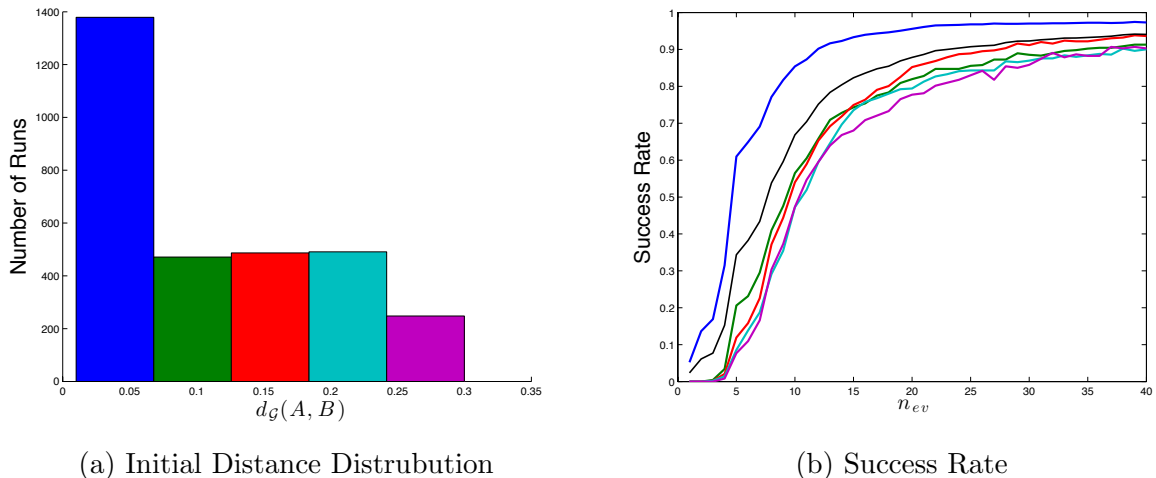
Figure 6.4: Success of the Race Car Algorithm for $n_{dof} = 6, n_{ev} = 10$. The initial and target shapes are the same as in Figure 6.3. The final shape distance is $d_{\mathcal{G}}(P(93), B) = 0.0008 < 0.01$. The final spectral distance is $d(\sigma(P(93)), \sigma(B)) = 5.61 \cdot 10^{-10} < 10^{-9}$.

(a) Initial Distance Distrubution       (b) Success Rate

Figure 6.5: Distribution of the sampled initial distances $d_{\mathcal{G}}(A, B)$ and success rate as a function of $n_{ev}$, from left to right. The colours of the lines of the second plot correspond to the colours of the bars in the first. The black curve is the overall success rate. Notice that the success rate increases with $n_{ev}$ and decreases with starting distance. Also note that at $n_{ev} = 5$, the qualitative behaviour of the plot changes. The success rate shows knowledge of the number of degrees of freedom.

By observing Figure 6.5, one immediately notices a few things. First, the success rate increases with $n_{ev}$, which is to be expected as more information is provided to the algorithm. Second, the success rate decreases with $d_{\mathcal{G}}(A, B)$, which is also intuitively plausible, since the longer the path the algorithm has to take from $A$ to $B$, the higher the chances to encounter a problematic point.

It is of immediate interest to explore how the success rate decays with increasing $d_{\mathcal{G}}(A, B)$. Notice that $d_{\mathcal{G}}(A, B) = 0$ implies automatic success. We can thus add such a point to our analysis of the dependence of the success rate on the distance. Figure 6.6 illustrates this dependence for fixed values of $n_{ev}$. Although the exact shape of the curve is coarse, Figure 6.6 suggests that for $n_{ev}$ fixed the dependence of the success rate upon $d_{\mathcal{G}}(A, B)$ is a decaying exponential of the form $\exp(-\rho(n_{ev})d_{\mathcal{G}}(A, B))$, where the decay rate $\rho(n_{ev})$ depends upon the number of eigenvalues considered by the algorithm. To determine the dependence of $\rho(n_{ev})$ upon $n_{ev}$, we use an exponential fit of the success rate as a function of $d_{\mathcal{G}}(A, B)$ for each value of $n_{ev}$. The obtained $\rho_{ev}$ are plotted on Figure 6.6. For $n_{ev} \geq 5$, $\rho(n_{ev})$ follows a power law behaviour.
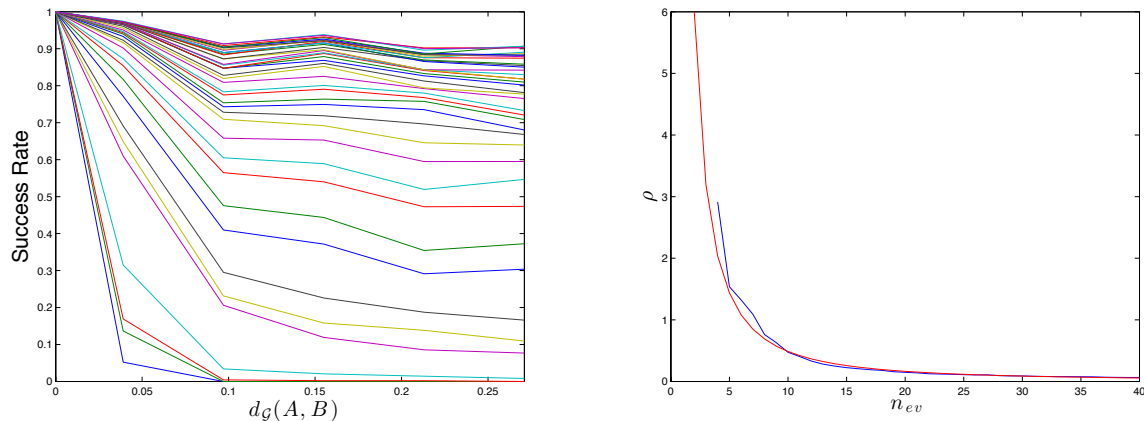
(a) Decay of the success rate with $d_{\mathcal{G}}(A, B)$      (b) Power law behaviour of $\rho(n_{ev})$

Figure 6.6: Approximately exponential decay of the success rate with $d_{\mathcal{G}}(A, B)$ and power law behaviour of the decay coefficient $\rho(n_{ev})$. On the left, each curve represents a given value of $n_{ev}$. Notice that the behaviour is consistent from one curve to the next. On the right, the blue curve represents the obtained values of $\rho(n_{ev})$ and the red curve is a power law fit of $\rho(n_{ev})$ for $n_{ev} \geq 5$. Exact values of $\rho(n_{ev})$ for $n_{ev} < 4$ could not be computed and thus are not included.

$$\rho(n_{ev}) \approx \alpha(n_{ev})^{-p} \quad , \quad n_{ev} \geq 5 \tag{6.29}$$

The constants $\alpha$ and $p$ of the power law behaviour of $\rho(n_{ev})$ are determined via a power law fit illustrated on Figure 6.6. The values are:

$$
\begin{aligned}
p &= 1.55 \pm 0.01 \\
\alpha &= 3.0 \pm 0.2
\end{aligned}
\tag{6.30}
$$

Assume that the failure of the algorithm is due to some sort of pathological points or regions in the space of shapes. Exponential decay of the success rate with $d_{\mathcal{G}}(A, B)$ suggests that the density of such obstacles is constant. In some fashion, $\rho(n_{ev})$ represents this density. Further assume that the pathological regions are essentially the same for all $n_{ev}$, such that, as one increases $n_{ev}$, some of the obstacles that the algorithm could not navigate with lower values of $n_{ev}$ become surmountable. That is, the pathological regions for $n_{ev} + 1$ are the same as for $n_{ev}$, minus a certain proportion that the algorithm now is

96

able to cross. The power law behaviour can then be exploited to determine how helpful it is to use more eigenvalues. Assume $n_{ev} \gg 1$.

$$\rho(n_{ev} + 1) = \rho\left(n_{ev}\left(\frac{n_{ev} + 1}{n_{ev}}\right)\right) = \left(\frac{n_{ev} + 1}{n_{ev}}\right)^{-p} \rho(n_{ev})$$
$$\approx \left(1 - \frac{p}{n_{ev}}\right) \rho(n_{ev}) \tag{6.31}$$

where the approximation follows from the Taylor series of the power law factor about 1. Assuming the above discussion is correct, we speculate that $\left(1 - \frac{p}{n_{ev}}\right)$ is the proportion of obstacles that remain insurmountable for the algorithm as $n_{ev} \rightarrow n_{ev} + 1$. Correspondingly, $p/n_{ev}$ is the proportion of obstacles that cease to be problematic. The decaying behaviour of $p/n_{ev}$ seems to indicate that higher eigenvalues contain less and less useful information for the reconstruction of the shape. It would be of great interest to explain this dependence. Plausibly, this would come with the understanding of the exact nature of the pathological regions in shape space. Further investigation is required.

A first step on the way to the understanding of the failure modes of the algorithm can be made by studying the number of isospectral non-isometric shapes found. Indeed, it is expected that there exist numerous shapes that share the first $n_{ev}$ eigenvalues of their spectrum. There is thus nothing mysterious with this type of failure. Recall that we set $\varepsilon_{\mathcal{G}} = 0.01$ and $\varepsilon_{\sigma} = 10^{-9}$. In order to restore coherence between the isospectrality and isometry thresholds, it is necessary to set $\varepsilon_{\sigma} = 5 \cdot 10^{-8}$. The dependence of the failure rate on $n_{ev}$ and $d_{\mathcal{G}}(A, B)$ is illustrated on Figure 6.7.

Figure 6.7 shows that about half of the observed failures can be attributed to $n_{ev}$-isospectral non-isometric shapes. It also seems to indicate that this proportion decays slowly as $n_{ev}$ increases. This is to be expected, as increasing the number of considered eigenvalues also has to increase the number of similar characteristics two shapes must have in order to be isospectral. Moreover, the proportion of isospectral failures decreases dramatically as $d_{\mathcal{G}}(A, B)$ increases. This can be explained by the fact that if $A$ is far away from $B$, it is likely that it is sufficiently different from $B$ not to have any shapes $n_{ev}$-isospectral to $B$ in its immediate neighbourhood. Thus, for some stretch of its path towards $\sigma(B)$ it can only encounter problematic points that are not isospectral to $B$. As mentioned earlier, the nature of such points will be the subject of future research. By the above discussion, studying cases where $d_{\mathcal{G}}(A, B)$ is large should be a reliable way of generating plenty of such points.
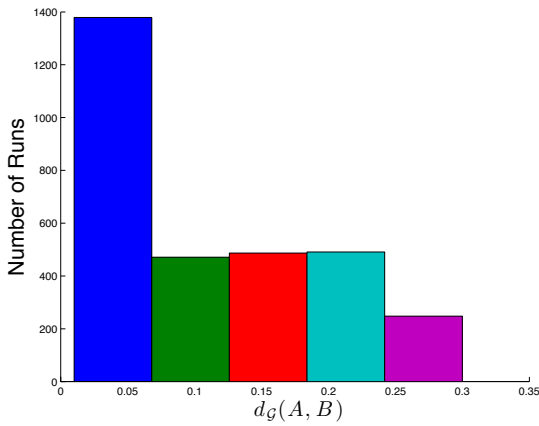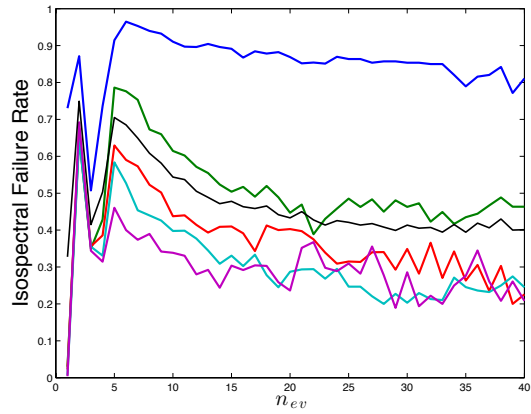
(a) Initial Distance Distrubution       (b) Isospectral Failure Rate

Figure 6.7: Distribution of the sampled initial distances $d_{\mathcal{G}}(A, B)$ and isospectral failure rate as a function of $n_{ev}$, from left to right. The colours of the lines of the second plot correspond to the colours of the bars in the first. The black curve is the overall isospectral failure rate. Notice that the isospectral failure rate slowly decreases with $n_{ev}$ and dramatically decreases with starting distance. Once again, the qualitative behaviour of the curve changes at $n_{ev} = 5$. Notice that about half of all observed failures are of the isospectral non-isometric type. For now, the precise functional dependence of the isospectral failure rate on $n_{ev}$ and $d_{\mathcal{G}}(A, B)$ cannot be determined, as the collected failure statistics are poor.

# Chapter 7

# Outlook

The numerical results presented in Chapter 6 are quite encouraging. Indeed, despite the volume of literature devoted to counterexamples in inverse spectral geometry, the simple implementation of the pseudoinverse method that we proposed can produce success rates on the order of 80 to 95% in the explored regime. It was also determined that, about half of the failure cases seem to be due to $n_{ev}$-isospectral non-isometric manifolds and that this ratio decayed with the number of considered eigenvalues. It is thus plausible that those problematic cases could be eliminated by studying more eigenvalues. However, the remaining failure cases are presently not understood.

A description of the failure cases of the pseudoinverse method would be of great importance, as it would suggest how to modify the algorithm in order to increase the success rate. As suggested in Chapter 6, non-isospectral failure cases can be generated by choosing the initial and target shapes to be very different, so that the algorithm first has to cross a long stretch of $\mathbb{R}^{n_{dof}}$ where it is unlikely to encounter shapes isospectral to the target one. Then, the only possible types of failure are of the non-isospectral kind. Once those problematic points or regions are found, one should first verify that they are not caused by numerical error. For instance, one could re-compute the Jacobian using a finer discretization of the manifold (i.e. a better approximation of the spectrum) and see if the situation improves. If so, no further explanation is needed. Otherwise, the problem is of fundamental nature and should be carefully explored. For example, as suggested to us by Kempf [46], one could look at the angle between $v_\sigma$, the direction in which the algorithm is supposed to go in spectral space, and the actual direction it goes in. Knowing this would allow one to determine if the obstacles are hit head on or tangentially. Moreover, it would be interesting to determine if there is a way to anticipate the encounter of problematic points. For instance, the effective rank could be a predictor of problematic points. Another possibility would be
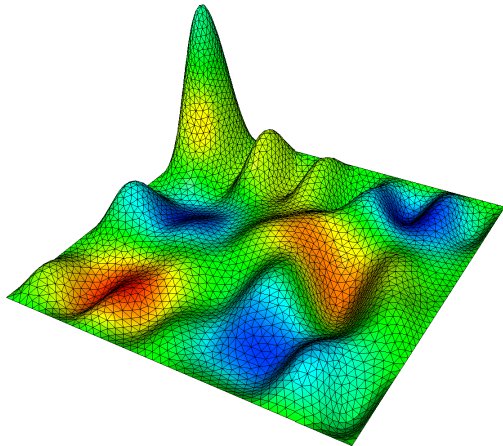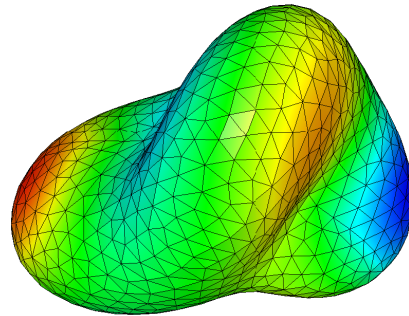
(a) Dirichlet boundary conditions         (b) Boundaryless

Figure 7.1: A taste of things to come. Two surfaces of dimension two embedded in $\mathbb{R}^3$, one with boundaries and one without. The colour scale represents the ninth eigenfunction, in both cases. Dirichlet boundary conditions are imposed in the case with boundaries. In that case, the shape is described by a two-dimensional Fourier sine series on the square $(0,1) \times (0,1)$. The boundaryless manifold is diffeomorphic to $S^2$ and its shape is described by a radius function given by the exponential of a series of spherical harmonics. An implementation of numerical infinitesimal inverse spectral geometry for such shapes is within reach.

to draw a map of the obstacles. This could be used to verify how their position depends upon the target shape $B$. Most interesting are of course those pathological points which do not depend upon the target, as they are fundamental obstacles to infinitesimal inverse spectral geometry.

It is also within immediate reach to apply the methods of infinitesimal inverse spectral geometry to different settings. In order to push the investigation closer to its physical motivations, it is necessary to study how it behaves on manifolds with curvature. The simplest possible examples of this would be two dimensional surfaces embedded in $\mathbb{R}^3$ equipped with the induced metric. Examples of such surfaces are illustrated on Figure 7.1. In such cases, the shape of the manifold is described as a height or radius function given by a combination of a finite number of known functions. This is entirely analogous to the

way the shape of a boundary was described in Chapter 6.

Possibly, one could also attempt to apply the numerical methods that we proposed to dynamical systems theoretic characterizations of shapes, such as billiard orbits or random walks on manifolds. This could shed a different light on infinitesimal inverse spectral geometry.

It could also be of interest to study some of the inverse spectral problems analogous to inverse spectral geometry using the proposed methods. For instance, spectra of Schrödinger operators in one dimension should be easier to compute than those of Laplacians on manifolds. Thus, that setting could be used to quickly test the performance of implementations of infinitesimal ISG.

A possible practical application of infinitesimal inverse spectral geometry could be the detection of the number of degrees of freedom of a system. Figures 6.5 and 6.7 show that the rates of success and isospectral failure of the pseudoinverse approach have a qualitatively different behaviour for $n_{ev} < 5$ and for $n_{ev} \geq 5$. Recall that while formally the shapes studied had 6 degrees of freedom, in effect one of them was simply a choice of orientation. Thus, those results indicate that infinitesimal inverse spectral geometry is sensitive to the actual number of degrees of freedom in a system. Abstractly, this could be of interest to determine how much redundancy there is in a poorly understood construction map $C : \mathbb{R}^{n_{dof}} \to \mathcal{G}$. Practically, it could be used to determine when a physical system has gained unwanted degrees of freedom. This could have engineering applications, as there unwanted degrees of freedom usually indicate wear and tear. Applications to quantum gravity will have to wait until the simpler cases are well understood.

# APPENDICES

# Appendix A

# Moore-Penrose Pseudoinverse

This appendix is a very brief introduction to the pseudoinverse of a linear map between finite-dimensional vector spaces, sometimes referred to as the Moore-Penrose pseudoinverse or Moore-Penrose generalized inverse. We refer the reader to [64] for a concise and clear exposition of this subject. We present an abridged version of the discussion therein.

Let $U, V$ be inner product spaces over $\mathbb{C}$ or $\mathbb{R}$ of finite dimensions $p$ and $q$, respectively. Let $L : U \to V$ be a linear transformation of rank $r$. For a number of reasons, $L$ might not be invertible. Nonetheless, there are many practical cases in which one would like to invert $L$, at least as much as possible. Let $L^\dagger$ denote the adjoint of $L$. It is possible to decompose $U = A \oplus B$ and $V = C \oplus D$ as direct sums of carefully chosen vector subspaces. Consider the following theorem.

**Theorem A.1.** *Let $L, L^\dagger$ be as above. The images and kernels of $L, L^\dagger$ and of their compositions have the following relationships*

*(a) $ker(L^\dagger) = im(L)^\perp$ and $im(L^\dagger) = ker(L)^\perp$*

*(b) $ker(L^\dagger L) = ker(L)$ and $ker(LL^\dagger) = ker(L^\dagger)$*

*(c) $im(L^\dagger L) = im(L^\dagger)$ and $im(LL^\dagger) = im(L)$*

*Proof.* See [64]. $\square$

$B$ is chosen to be the kernel of $L$ and $D$ is chosen to be the kernel of $L^\dagger$. One can introduce an orthonormal basis $\{u_i\}_{i=1}^r$ of $A$ and a basis $\{v_i\}_{i=1}^r$ of $C$ such that $L$ and $L^\dagger$ have the following symmetric action

$$Lu_i = s_i v_i$$
$$L^\dagger v_i = s_i u_i \tag{A.1}$$

This is guaranteed by the well known notion of *singular value decomposition*, which is the appropriate generalization of the eigenvalue decomposition for linear maps between different vector spaces. The $s_i$ of the above equation are called the *singular values* of $L$ and are defined below. Recall that one can define a non-negative self-adjoint operator $L^\dagger L : U \to U$. Its eigenvalues can be ordered as follows

$$\lambda_1 \geq ... \geq \lambda_r > 0 = \lambda_{r+1} = ... = \lambda_p \tag{A.2}$$

The corresponding orthonormal eigenbasis is given by

$$\mathcal{B} = (u_1, ..., u_r, u_{r+1}, ..., u_n) \tag{A.3}$$

Note that $(u_{r+1}, ..., u_p)$ is a basis for $ker(L^\dagger L) = ker(L)$. Necessarily, $(u_1, ..., u_r)$ is a basis for $ker(L)^\perp = im(L^\dagger)$. The singular values of $L$ are defined to be the positive square roots of the nonzero eigenvalues of $L^\dagger L$

$$s_i = \sqrt{\lambda_i} \ , \quad i = 1...r \tag{A.4}$$

We now choose a particular orthonormal basis of $im(L)$:

$$v_i = \frac{1}{s_i} L u_i \ , \quad i = 1...r \tag{A.5}$$

The proof of the orthonormality of this basis is straightforward and is left to the reader. This basis can be augmented by vectors $(v_{r+1}, ..., v_q)$ to form an orthonormal basis for all of $V$. In the particular bases chosen, the action of $L$ and $L^\dagger$ takes the promised symmetric form:

$$Lu_i = \begin{cases} s_i v_i & i \leq r \\ 0 & i > r \end{cases} \quad ; \quad L^\dagger v_i = \begin{cases} s_i u_i & i \leq r \\ 0 & i > r \end{cases} \tag{A.6}$$

The above basis is used to define the Moore-Penrose pseudoinverse $L^+$ of $L$.

$$L^+ v_i := \begin{cases} \frac{1}{s_i} u_i & i \le r \\ 0 & i > r \end{cases} \tag{A.7}$$

When the pseudoinverse is restricted to act between the subspaces $A$ and $C$, it indeed reproduces the inverse

$$\begin{aligned} (L^+ L)\big|_A &= \mathbb{1}_A \\ (L L^+)\big|_C &= \mathbb{1}_C \end{aligned} \tag{A.8}$$

When restricted to act on the kernels of $L$ and $L^\dagger$, the pseudoinverse acts trivially, as required

$$\begin{aligned} (L^+ L)\big|_B &= 0 \\ (L L^+)\big|_D &= 0 \end{aligned} \tag{A.9}$$

If $L$ is invertible, $L^+ = L^{-1}$. Note that, by construction, $ker(L^+) = ker(L^\dagger)$ and $im(L^+) = im(L^\dagger)$.

It is possible to define the pseudoinverse by algebraic means, without ever picking particular bases for $U$ and $V$.

**Theorem A.2.** *Let L be as above. Its pseudoinverse $L^+$ is the unique operator characterized by*

*(a)* $LL^+ L = L$

*(b)* $L^+ L L^+ = L^+$

*(c)* $LL^+$ *is self-adjoint*

*(d)* $L^+ L$ *is self-adjoint*

*This definition is equivalent to the constructive one given prior.*

*Proof.* See [64]. □

The pseudoinverse is used to solve *least squares problems*. Let $Lx = v$ represent a system of linear equations, with $x$ being the unknown. This system has a solution if and only if $v \in im(L)$. Assume that the system of equations has no solution. Let $\hat{v} \in im(L)$ denote a vector such that $\|v - \hat{v}\|_2$ is minimal. This vector can be shown to be unique. One can replace the initial problem with $Lx = \hat{v}$. Any $x$ satisfying that system of equations is said to be a solution to the linear least squares problem. In terms of $v$, $x$ is such that $\|Lx - v\|_2$ is minimal.

In general, there are infinitely many possible values of $x$. Consider $x$ and $x'$, both solutions to the least squares problem. Since $Lx = \hat{v} = Lx'$, $x - x' \in ker(L)$. Consequently, given a least squares solution $x$, all other least squares solutions are given by $x + ker(L)$. If the dimension of $ker(L)$ is nonzero, least squares solutions can have arbitrarily large norms. Thus, it is particularly interesting to determine the solution of least norm, which is easily shown to be unique. Essentially, this amounts to set the components of $x$ in the kernel to zero. This can be achieved by setting $x = L^+v$. See [64] for a proof. Thus, any application of the pseudoinverse can be seen as a solution to a particular least squares problem.

# Appendix B

# An Absolutely Minimal Introduction to Finite Element Methods

Finite element methods provide powerful ways to numerically solve partial differential equations. This is a field rich in both mathematical insight and practical applications. The goal of this appendix is to provide the reader an *absolute minimum* of understanding of those methods. Expect no mathematical rigour in the below discussion. The interested reader may consult [9] for a more thorough introduction. An intriguing discussion on the usage of topological and homological techniques in finite element methods can be found in [2].

We will introduce finite elements with the help of an example shown to us by Helmut Kröger [49]. Consider the following first order boundary value problem:

$$-\frac{d}{dx}y(x) + y(x) - 1 = 0 \tag{B.1}$$
$$y(0) = y(\pi) = 0$$

The analytic solution for this problem is given by

$$y(x) = \frac{e^x - e^{\pi - x}}{e^\pi + 1} - 1 \tag{B.2}$$

For the sake of notational simplicity, set $L = -\frac{d}{dx} + 1$. Suppose that we are to find this solution by a numerical approach. We first extend the problem from the space of differentiable functions to $L_2(0, \pi)$. Let $\{\psi_i\}_{i=1}^{\infty}$ be a basis of this Hilbert space, not necessarily orthonormal. It is thus possible to express $y(x)$ as

$$y(x) = \sum_{i=1}^{\infty} y_i \psi_i(x) \tag{B.3}$$

Suppose that the first $n$ basis functions are zero for $x = 0, \pi$, which ensures that all of their linear combinations satisfy the boundary conditions. Moreover, suppose that the first $n$ basis functions offer an acceptable approximation of $y(x)$. Let $\tilde{y}(x)$ be such an approximation of $y(x)$.

$$\tilde{y}(x) = \sum_{i=1}^{n} y_i \psi_i(x) \tag{B.4}$$

It is expected that $\tilde{y}(x)$ will not exactly satisfy the boundary value problem. Let $r(x)$ be a residual term such that

$$L\tilde{y}(x) - 1 = r(x) \tag{B.5}$$

It is natural to require for $r(x)$ to be orthogonal to the first $n$ basis functions.

$$(\psi_i(x), r(x))_{L_2} = 0 \quad, \quad i = 1...n$$
$$\sum_{i=1}^{n} y_i(\psi_j(x), \psi_i(x))_{L_2} + (\psi_j, 1)_{L_2} = 0 \tag{B.6}$$

The above can be recast in matrix form by setting $M_{ji} = (\psi_j(x), \psi_i(x))_{L_2}$ and $f_j = (\psi_j, 1)_{L_2}$. Solving the approximate problem reduces to solving the following system of linear equations in matrix form.

$$\sum_{i=1}^{n} M_{ji} y_i = f_i \tag{B.7}$$

The problem has thus been moved entirely within the realm of linear algebra. The usage of a truncated basis that satisfies boundary conditions is known as Galerkin's method. The above is clearly generalizable to other linear operators.

The passage from Galerkin's method to finite elements is a choice of basis. Consider for instance a basis of tent functions as in figure B.1. More generally, one would choose a basis of (almost everywhere) sufficiently differentiable functions with *overlapping compact*
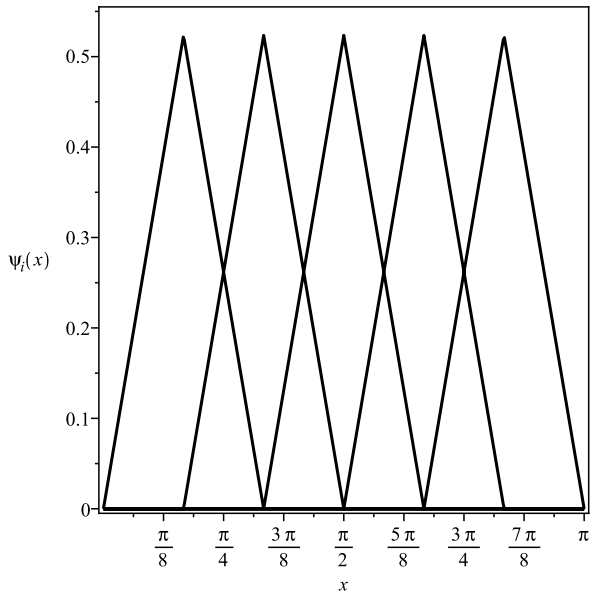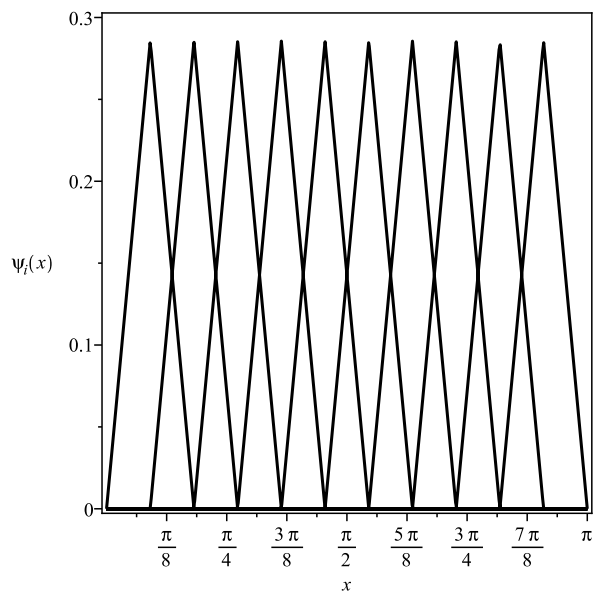
(a) Tent basis for $n = 5$    (b) Tent basis for $n = 10$

Figure B.1: A simple finite element basis of piecewise linear functions. On the left, five basis functions are used, while on the right, ten are.

*supports.* Typically, one chooses piecewise polynomial functions. The finiteness of the size of the support of the basis functions is what makes the elements finite. Figure B.2 compares the exact solution to the problem posed by Equation (B.1) to finite element solutions obtained with the tent function bases of Figure B.1. Notice that the approximate solutions are in very good agreement with the exact one.
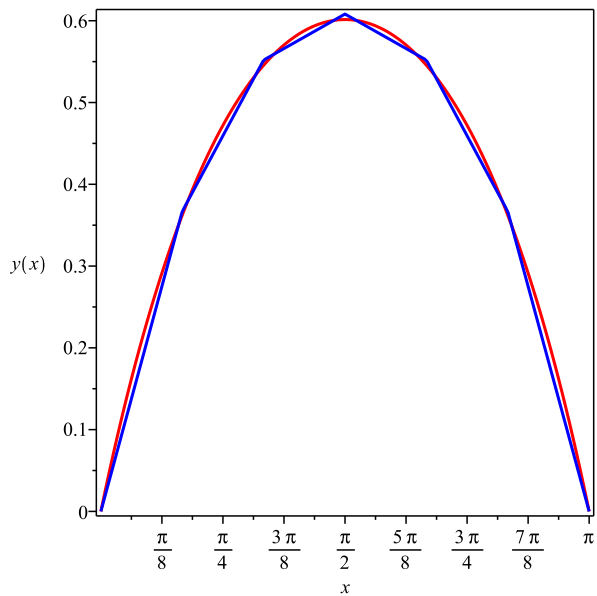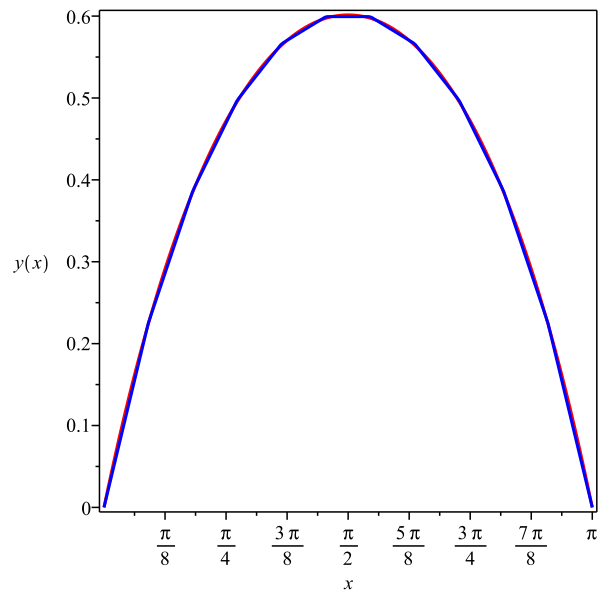
(a) $n = 5$             (b) $n = 10$

Figure B.2: Comparison between exact and finite element solutions to the problem defined in Equation (B.1). The red curve is the exact solution given by equation (B.2). The blue curves are finite element approximations of it, using the tent function bases introduced in Figure B.1. The agreement between the exact and approximate solutions is excellent.

# References

[1] D. Aasen, T. Bhamre, and A. Kempf. Shape from sound: toward new tools for quantum gravity. *Physical Review Letters*, 110(12):121301, 2013.

[2] D. N. Arnold, R. S. Falk, and R. Winther. Finite element exterior calculus, homological techniques, and applications. *Acta numerica*, 15(1):1–155, 2006.

[3] S. Bando and H. Urakawa. Generic properties of the eigenvalue of the Laplacian for compact Riemannian manifolds. *Tohoku Mathematical Journal*, 35(2):155–172, 1983.

[4] P. H. Bérard and G. Besson. *Spectral geometry: direct and inverse problems*, volume 1207. Springer-Verlag Berlin, 1986.

[5] M. Berger. *A panoramic view of Riemannian geometry*. Springer, 2003.

[6] N. D. Birrell and P. C. W. Davies. *Quantum fields in curved space*. Cambridge University Press, 1984.

[7] E. d. S. L. Brandao. On infinitesimal inverse spectral geometry. Master's thesis, University of Waterloo, 2011.

[8] T. P. Branson and P. B. Gilkey. The asymptotics of the Laplacian on a manifold with boundary. *Communications in partial differential equations*, 15(2):245–272, 1990.

[9] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 2008.

[10] H. Brezis. *Functional analysis, Sobolev spaces and partial differential equations*. Springer, Breiningsville, PA, 2010.

[11] R. Brooks. The Sunada method. *Contemporary Mathematics*, 231:25–36, 1999.

[12] S. R. Buss. Introduction to inverse kinematics with Jacobian transpose, pseudoinverse and damped least squares methods. *IEEE Journal of Robotics and Automation*, 17, 2004.

[13] I. Chavel. *Eigenvalues in Riemannian geometry*. Academic Press, Orlando, FL, 1984.

[14] F. R. Chung. *Spectral graph theory*, volume 92 of *CBMS Regional Conference Series in Mathematics*. AMS, 1997.

[15] A. Connes. *Noncommutative geometry*. Academic press, 1995.

[16] C. B. Croke and V. A. Sharafutdinov. Spectral rigidity of a compact negatively curved manifold. *Topology*, 37(6):1265–1273, 1998.

[17] K. Datchev and H. Hezari. Inverse problems in spectral geometry. *Inverse Problems and Applications: Inside Out II*, 2011.

[18] J. Dodziuk. Eigenvalues of the Laplacian and the heat equation. *The American Mathematical Monthly*, 88(9):686–695, 1981.

[19] M. S. Elnaggar and A. Kempf. On a generic entropy measure in physics and information. In *Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009. WiOPT 2009. 7th International Symposium on*, pages 1–5. IEEE, 2009.

[20] M. S. Elnaggar, S. Safavi-Naeini, and S. K. Chaudhuri. A novel dimensionality metric for multi-antenna systems. In *Microwave Conference, 2006. APMC 2006. Asia-Pacific*, pages 242–245. IEEE, 2006.

[21] H. Federer. *Geometric Measure Theory*. Springer-Verlag, 1969.

[22] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, 1966.

[23] F. Gesztesy and B. Simon. On the determination of a potential from three spectra. In *Differential operators and spectral theory*, volume 189, pages 85–92. American Mathematical Society Translations, AMS, 1999.

[24] P. Gilkey, J. Leahy, and J. Park. Eigenvalues of the form valued Laplacian for Riemannian submersions. *Proceedings of the American Mathematical Society*, 126(6):1845–1850, 1998.

[25] P. B. Gilkey. *Invariance theory, the heat equation, and the Atiyah-Singer index theorem.* CRC Press, Boca Raton, FL, second edition, 1994.

[26] P. B. Gilkey. *Asymptotic formulae in spectral geometry.* CRC press, 2003.

[27] C. Gordon, D. L. Webb, and S. Wolpert. One cannot hear the shape of a drum. *Bulletin of the American Mathematical Society*, 27:134–138, 1992.

[28] C. Gordon and E. Wilson. Isospectral deformations of compact solvmanifolds. *Journal of Differential Geometry*, 19(1084):241–256, 1984.

[29] C. S. Gordon. Survey of isospectral manifolds. *Handbook of differential geometry*, 1:747–778, 2000.

[30] M. D. Greenberg. *Advanced Engineering Mathematics.* Prentice-Hall, second edition, 1998.

[31] P. Greiner. An asymptotic expansion for the heat equation. *Archive for Rational Mechanics and Analysis*, 41(3):163–218, 1971.

[32] G. Grubb. *Functional calculus of pseudo-differential boundary problems*, volume 65 of *Progress in Mathematics.* Birkhauser, second edition, 1996.

[33] V. Guillemin and D. Kazhdan. Some inverse spectral results for negatively curved 2-manifolds. *Topology*, 19(3):301–312, 1980.

[34] V. Guillemin and R. Melrose. An inverse spectral result for elliptical regions in $\mathbb{R}^2$. *Advances in Mathematics*, 32(2):128–148, 1979.

[35] D. Gurarie. Semiclassical eigenvalues and shape problems on surfaces of revolution. *Journal of Mathematical Physics*, 36:1934, 1995.

[36] A. Hassell and S. Zelditch. Determinants of Laplacians in exterior domains. *International Mathematics Research Notices*, 1999(18):971–1004, 1999.

[37] M. Hein, J.-Y. Audibert, and U. Von Luxburg. From graphs to manifolds–weak and strong pointwise consistency of graph Laplacians. In *Learning theory*, pages 470–485. Springer, 2005.

[38] D. Henry, J. K. Hale, and A. L. Pereira. *Perturbation of the boundary in boundary-value problems of partial differential equations*, volume 318 of *London Mathematical Society Lecture Note Series.* Cambridge University Press, 2005.

[39] H. Hezari and S. Zelditch. $C^\infty$ spectral rigidity of the ellipse. *preprint*, 2010.

[40] H. Hezari and S. Zelditch. Inverse spectral problem for analytic $(\mathbb{Z}/2\mathbb{Z})$-symmetric domains in $\mathbb{R}^N$. *Geometric and Functional Analysis*, 20(1):160–191, 2010.

[41] D. Husemöller. *Fibre bundles.* Springer, third edition, 1994.

[42] K. Itō. *Encyclopedic Dictionary of Mathematics.* The MIT Press, second edition, 1996.

[43] D. Jakobson, M. Levitin, N. Nadirashvili, and I. Polterovich. Spectral problems with mixed Dirichlet–Neumann boundary conditions: isospectrality and beyond. *Journal of computational and applied mathematics*, 194(1):141–155, 2006.

[44] J. Jost. *Riemannian geometry and geometric analysis.* Springer, fifth edition, 2008.

[45] M. Kac. Can one hear the shape of a drum? *The American Mathematical Monthly*, 73(4, Part 2: Papers in Analysis):1–23, 1966.

[46] A. Kempf. personal communication.

[47] A. Kempf. Spacetime could be simultaneously continuous and discrete, in the same way that information can be. *New Journal of Physics*, 12(11):115001, 2010.

[48] A. Kempf. Quantum gravity on a quantum computer? *Foundations of Physics*, 2013.

[49] H. Kröger. personal communication, 2010.

[50] R. Kuwabara. On the characterization of flat metrics by the spectrum. *Commentarii Mathematici Helvetici*, 55(1):427–444, 1980.

[51] G. Landi and C. Rovelli. General relativity in terms of Dirac eigenvalues. *Physical review letters*, 78(16):3051, 1997.

[52] H. P. McKean and E. Trubowitz. The spectral class of the quantum-mechanical harmonic oscillator. *Communications in Mathematical Physics*, 82(4):471–495, 1982.

[53] R. Melrose. Isospectral sets of drumheads are compact in $C^\infty$. preprint, 1983.

[54] B. L.-M. Michelson and H. Lawson. *Spin geometry.* Princeton, 1989.

[55] J. Milnor. Eigenvalues of the Laplace operator on certain manifolds. *Proceedings of the National Academy of Sciences of the United States of America*, 51(4):542, 1964.

[56] B. Mohar and Y. Alavi. The Laplacian spectrum of graphs. *Graph theory, combinatorics, and applications*, 2:871–898, 1991.

[57] A. W. Naylor and G. R. Sell. *Linear operator theory in engineering and science*. Springer, 2000.

[58] J. Nestruev. *Smooth manifolds and observables*. Springer, 2003.

[59] L. I. Nicolaescu. *Lectures on the geometry of manifolds*. World Scientific Publishing, Toh Tuck Link, Singapore, second edition, 2007.

[60] V. Patodi. Curvature and the fundamental solution of the heat operator. *Collected Papers of VK Patodi*, 34:51, 1996.

[61] V. M. Petkov and L. N. Stoyanov. *Geometry of Reflecting Rays and Inverse Spectral Problems*. John Wiley and Sons Inc., 1992.

[62] V. N. Pivovarchik. An inverse Sturm-Liouville problem by three spectra. *Integral Equations and Operator Theory*, 34(2):234–243, 1999.

[63] N. Qian. On the momentum term in gradient descent learning algorithms. *Neural networks*, 12(1):145–151, 1999.

[64] S. Roman. *Advanced linear algebra*, volume 135. Springer, third edition, 2008.

[65] O. Roy and M. Vetterli. The effective rank: A measure of effective dimensionality. *Entropy*, 4:7, 2007.

[66] B.-W. Schulze. *Pseudo-differential operators on manifolds with singularities*, volume 24 of *Studies in Mathematics and its Applications*. Elsevier, 1991.

[67] B.-W. Schulze. *Boundary Value Problems and Singular Pseudo-Differential Operators*. Pure and Applied Mathematics. Wiley, 1998.

[68] V. A. Sharafutdinov. Local audibility of a hyperbolic metric. *Siberian Mathematical Journal*, 50(5):929–944, 2009.

[69] M. A. Shubin. *Pseudodifferential Operators and Spectral Theory*. Springer, second edition, 2001.

[70] T. Sunada. Riemannian coverings and isospectral manifolds. *Ann. of Math.(2)*, 121(1):169–186, 1985.

[71] Z. I. Szabo. Locally non-isometric yet super isospectral spaces. *Geometric And Functional Analysis*, 9(1):185–214, 1999.

[72] P. Szekeres. *A course in modern mathematical physics: groups, Hilbert spaces and differential geometry*. Cambridge University Press, New York, 2004.

[73] S. Tanno. Eigenvalues of the Laplacian of Riemannian manifolds. *Tôhoku Mathematical Journal*, 25(3):391–403, 1973.

[74] S. Tanno. A characterization of the canonical spheres by the spectrum. *Mathematische Zeitschrift*, 175(3):267–274, 1980.

[75] S. Zaremba. Sur un problème toujours possible comprenant, à titre de cas particulier, le problème de Dirichlet et celui de Neumann. *Journal de Mathématiques pures et appliquées*, pages 127–164, 1927.

[76] S. Zelditch. The inverse spectral problem for surfaces of revolution. *Journal of Differential Geometry*, 49(2):207–264, 1998.

[77] S. Zelditch. Inverse spectral problem for analytic plane domains II: $\mathbb{Z}_2$-symmetric domains. *arXiv preprint math/0111078*, 2001.

[78] G. Zhou. Compactness of isospectral compact manifolds with bounded curvatures. *Pacific Journal of Mathematics*, 181(1), 1997.