KERNEL-BASED GALERKIN METHODS ON COMPACT MANIFOLDS WITHOUT BOUNDARY, WITH AN EMPHASIS ON $SO(3)$

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAIʻI AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

MATHEMATICS

August 2021

By

Patrick Collins

Dissertation Committee:

Thomas Hangelbroek, Chairperson
Wayne Smith
Rufus Willett
Evan Gawlik
Marcelo Kobayashi

Abstract

We develop a kernel-based Galerkin method for numerically solving elliptic partial differential equations on compact Riemannian manifolds without boundary that are *mesh-free*, *coordinate-free*, *regularity-preserving*, and *high-performing*. A fundamental challenge is to combine a theoretical solution with mesh-free *quadrature* in such a way that approximation power is not lost.

We show that the approximation power of the computed (discretized) solution can be made to be on par with the approximation power of the theoretical solution, provided the *oversampling exponent* is sufficiently large. We then show how to truncate a kernel on $SO(3)$ given by a Hilbert-Schmidt series in such a way that the approximation power of both the truncated solution and the discretized truncated solution (again using quadrature) is on par with the approximation power of the theoretical solution, provided the *truncation parameter* is sufficiently large.

Contents

List of Figures

Chapter 1

Introduction

Our overarching goal is to develop kernel-based methods for numerically solving elliptic partial differential equations (PDEs) on compact Riemannian manifolds that are *mesh-free*, *coordinate-free*, *high-performing*, and *regularity-preserving*. Mesh-free means we don't need to build any additional structure on our underlying space - like triangulations or uniform grids - only to sample the functions that appear in the differential equation. Coordinate-free means when we are considering a problem on a manifold, we don't have to work in coordinate patches, and therefore don't have to join together pieces of solutions. We note that we will work in coordinate patches to prove various results, but the methods themselves don't require it. High-performing means the methods work progressively better for more regular solutions, and can thus be suited to the regularity of the solutions we want to approximate. Regularity-preserving means that the regularity of the computed solution matches the regularity of the actual solution.

A Galerkin method is one in which the equation is put in a weak form, which is an equation on a Hilbert space with a bilinear form on the left and a linear functional on the right. Putting the equation in weak form usually requires multiplying by a *test function* and performing a manipulation analogous to integration by parts. The corresponding weak form of the problem is then solved on a finite-dimensional subspace of that Hilbert space.

We demonstrate with a classical example on the circle. Consider the second-order equation $-u'' + u = f$ on $\mathbb{S}^1$, where $f \in L_2(\mathbb{S}^1)$. Our set of test functions in this context is $H^1(\mathbb{S}^1)$, the set of functions $v \in L_2(\mathbb{S}^1)$ for which the (distributional) derivative v' is also in $L_2(\mathbb{S}^1)$. Multiplying by $v \in H^1(\mathbb{S}^1)$, integrating, and performing integration by parts on the principal part on the left yields our weak form:

$$\langle u, v\rangle := \int_0^{2\pi} \left(u'(\theta)v'(\theta) + u(\theta)v(\theta)d\theta = \int_0^{2\pi} f(\theta)v(\theta) \right) d\theta =: \lambda_f(v).$$

The weak formulation of the problem now reads: find $u \in H^1(\mathbb{S}^1)$ such that $\langle u, v\rangle = \lambda_f(v)$ for all $v \in H^1(\mathbb{S}^1)$. The finite-dimensional subspace V of $H^1(\mathbb{S}^1)$ that we use could be, for example, a space of trigonometric polynomials, or a suitable space of piecewise-defined algebraic polynomials. A *Galerkin approximation* is a solution to the following modified problem: find $u \in V$ such that $\langle u, v\rangle = \lambda_f(v)$ for all $v \in V$. Given a basis $\mathcal{B}$ for V, the coefficients of the Galerkin approximation in that basis are usually obtained by solving an $n \times n$ linear system, where $n = \dim(V)$ - for each $b \in \mathcal{B}$, $\langle u, b\rangle = \lambda_f(b)$ yields an equation in the n unknown coefficients, and there are n such b. The matrix of this linear system is called the *stiffness matrix.*

We consider in this thesis kernel-based Galerkin methods. Other prominent Galerkin methods include *finite-element methods*, where the finite-dimensional spaces are functions supported in subdomains of the underlying space - for example, piecewise polynomials on a triangulation. Methods in which trigonometric polynomials or, more generally, eigenfunctions of the Laplace-Beltrami operator, are often referred to as *spectral methods*. Indeed, the truncated Galerkin approximation scheme we treat in Chapter 6 is a spectral method.

In order to implement these methods practically we must find a way to suitably discretize the bilinear form and linear functional. In our case we use *quadrature*, in which integrals are approximated via a weighted average of samples of the integrand, analogous to Simpson's rule or the trapezoid rule in Calculus. This poses a significant challenge. Analyzing the error for the problem using the true stiffness matrix is fairly elementary (we provide this error analysis in Section 3.3). Modifying the problem to treat a nearby stiffness matrix whose entries are obtained with quadrature is much more difficult. We show that, on certain manifolds and with certain kernels, there is a method that is implementable and provides error estimates that can be made to be on par with those from the theoretical setting, where the bilinear form and linear functional are computed analytically.

Initially, the goal was to do this on the rotation group, $SO(3)$. We used as an inspiration the work done by Narcowich, Rowe, and Ward in [22], in which they develop a kernel-based Galerkin method on the sphere, $\mathbb{S}^2$. The main problem here for us is that the finite-dimensional spaces which occur naturally on $SO(3)$ are much more poorly localized than those on $\mathbb{S}^2$ considered in [22]. Some of our techniques work in a more general setting - on compact manifolds without boundary. Understanding how to do this on $SO(3)$ yields a way to do it on manifolds in the more general class, since all such spaces have natural kernels defined on them that generate bases with that kind of sub-optimal decay (but that decay nonetheless).

One special feature of the problem on $\mathbb{S}^2$ and $SO(3)$ is that there exist highly-performing kernels which can be given in closed form. For example, on the sphere,

$$\Phi(x, y) = \|x - y\|_2 \log \|x - y\|_2,$$

and on the rotation group,

$$\Phi(x, y) = \sin\left(\frac{\omega\left(y^{-1}x\right)}{2}\right)^{3/2},$$

where $\omega(x)$ gives the rotational angle of x. It may be a challenge on general manifolds to find closed-form formulas for kernels. In Chapter 6 we address this problem by developing a way to truncate highly-performing kernels that are given by a Hilbert-Schmidt series expansion. This gives a closed-form kernel, which we then show has the same approximation power as the original kernel, provided sufficiently many terms are included. Incidentally, our Galerkin method after we truncate is, technically, a spectral method.

We now discuss two major points in this thesis. The first, found in Chapter 4, is made in the general context, and says, roughly speaking, that the numerical approximation we get from quadrature has approximation power that matches the approximation power of the theoretical approximation, provided the samples used for quadrature are sufficiently denser than the samples used for Galerkin approximation. The second, found in Chapter 6, is made in the specific context of the rotation group, $SO(3)$, and says, roughly speaking, that the numerical approximation we get from truncation has approximation power that matches the approximation power of the theoretical approximation, provided the truncation parameter is large enough. We also show in Chapter 6 that, provided the truncation parameter is large enough *and* the samples used for quadrature are sufficiently denser than the samples used for Galerkin approximation, that the quadratized truncated Galerkin approximation has approximation power that matches the approximation power of the theoretical approximation. (We have coined the term "quadratized" because saying "discretized stiffness matrix" didn't seem right - all matrices are discrete.)

In Chapter 2, we discuss kernels, smoothness spaces (underlying Hilbert spaces), and present results on kernel-based interpolation. These are the classical problems considered in approximation theory; the results go back to [8].

In Chapter 3 we continue with some more contemporary kernel-based approximation results. We discuss quadrature, energy estimates, and exhibit a class of kernels - the polyharmonic kernels - that provide energy estimates on a certain class of manifolds - the two-point homogeneous manifolds. We end the chapter with an overview of the Galerkin method, including the theoretical results one gets without quadrature.

Our first main results appear in Chapter 4, as already mentioned. We establish error estimates between the weak solution to our equation, the Galerkin approximation, and the quadratized Galerkin approximation.

Chapter 5 applies the results of Chapter 4 to $SO(3)$ while developing two key examples. One, the rotational surface splines, has a closed form but only provides an algebraic energy estimate. The other provides an exponential energy estimate but doesn't have a closed form. The rotational surface splines are ready to use, whereas the one with no closed form has no usable formula.

In Chapter 6, we show how to work with the kernel from Chapter 5 that has no usable formula. We do this by truncating a Hilbert-Schmidt series for the kernel.

Our final Chapter **??** gives a brief discussion of how the work in this thesis could be continued.

At the beginning we assume only that our underlying space is a compact topological space and our kernel is either positive definite or conditionally positive definite. Assumptions on both the space and the kernel are added as they are required, until we have in chapter 4 a kernel that provides an energy estimate on a compact manifold without boundary. Finally, in chapter 6 we will be working with a specific kernel on a specific space, $SO(3)$.

Chapter 2

Background

In this chapter we lay the groundwork for subsequent chapters. In section 2.1, we introduce positive definite and conditionally positive definite kernels, and give some examples. We then explore the native spaces for these kinds of kernels in section 2.2. In section 2.3 we introduce our first notion of a Sobolev space: the H^τ spaces. We then discuss in section 2.4 our interpolation scheme and establish interpolation error estimates and some norm-minimizing properties of certain orthogonal projections onto our approximation spaces. While this thesis is not on differential geometry per se, it has some basis in geometry; the geometric background we need is laid out in section 2.5. In section 2.6 we introduce our second notion of a Sobolev space: the W_2^m spaces, introduce a local metric equivalence between our manifold and its tangent space, and prove a norm equivalence between H^m and W_2^m when m is a nonnegative integer. Section 2.7 defines various key quantities for our sets of centers and establishes the Zeros Lemma, a crucial result. We end the chapter with section 2.8, in which we state and prove an interpolation error estimate.

2.1 Positive Definite and Conditionally Positive Definite Kernels

Let Ω be a compact topological space. A *positive definite kernel* on Ω is a continuous function $\Phi : \Omega \times \Omega \to \mathbb{R}$ such that for every finite subset Ξ of Ω and every nonzero vector $\alpha = \{\alpha_\xi\}_{\xi \in \Xi} \in \mathbb{R}^\Xi$,

$$\sum_{\xi \in \Xi} \sum_{\eta \in \Xi} \alpha_\xi \alpha_\eta \Phi(\xi, \eta) > 0. \tag{2.1.1}$$

Equivalently, Φ is positive definite if and only if for every finite subset Ξ of Ω, the *collocation matrix*

$$K_\Xi = \{\Phi(\xi, \eta)\}_{\xi, \eta \in \Xi}$$

is a positive definite matrix. If an inequality as in (2.1.1) holds that is not strict, the kernel Φ is referred to as positive *semi*-definite.

If Π is a finite-dimensional subspace of $C(\Omega)$, then Φ is called *conditionally positive definite* with respect to Π on Ω if for every finite Π-unisolvent subset Ξ of Ω and every nonzero vector $\alpha \in \mathbb{R}^\Xi$ for which

$$\sum_{\xi \in \Xi} \alpha_\xi p(\xi) = 0$$

for all $p \in \Pi$, (2.1.1) holds. Being Π-unisolvent means that the only element p of Π for which $p(\xi) = 0$ for all $\xi \in \Xi$ is $p \equiv 0$. In this case the subspace Π is often referred to as the *auxiliary space*. So, Φ is conditionally positive definite with respect to Π on Ω if for every finite Π-unisolvent subset Ξ of Ω, (2.1.1) holds for all vectors $\alpha \in \mathbb{R}^\Xi$ that annihilate $\Pi|_\Xi$. The corresponding positive definite matrix here is the *augmented collocation matrix*

$$K_{\Xi,\Pi} = \begin{bmatrix} K_\Xi & P \\ P^T & 0 \end{bmatrix},$$

where $P = \{p_k(\xi)\}_{\xi\in\Xi,k=1,\ldots,Q}$, with $Q = \dim(\Pi)$ and $\{p_1,\ldots,p_Q\}$ a basis for Π. So, Φ is conditionally positive definite with respect to Π on Ω if for every finite Π-unisolvent subset Ξ of Ω, the augmented collocation matrix $K_{\Xi,\Pi}$ is a positive definite matrix. As before, if for all α that annihilate $\Pi|_\Xi$ an inequality like (2.1.1) holds that is not strict, Φ is referred to as conditionally positive *semi*-definite.

Here are some examples of positive definite and conditionally positive definite kernels in the Euclidean setting.

Example 2.1.2. (a) The *Gaussians* $\Phi(x,y) = e^{-a\|x-y\|_2^2}$, with $a > 0$, are positive definite on $\mathbb{R}^d$ for all d.

(b) If ϕ is continuous and in $L_1\left(\mathbb{R}^d\right)$, then $\Phi(x,y) = \phi(x-y)$ is positive definite on $\mathbb{R}^d$ if the Fourier transform $\widehat{\phi}$ of ϕ is nonnegative and non-vanishing. This is due to Bochner's Theorem - see [29, Section 6.2] for details.

(c) $\Phi(x,y) = \|x-y\|_2^2\log\|x-y\|_2$ is conditionally positive definite on $\mathbb{R}^2$. It is, modulo a multiplicative constant, the fundamental solution to Δ^2 on $\mathbb{R}^2$. (The operator Δ here is the usual Laplacian, $\frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2}$.) This means that for $f \in C^4\left(\mathbb{R}^2\right)$, we have the reproduction $f(x) = \int_{\mathbb{R}^2} \Phi(x,y)\Delta_y^2 f(y)dy$ for all $x \in \mathbb{R}^2$. The auxiliary space here is Π_1, the space of bivariate polynomials of degree at most 1. This is one of the kernels treated by Duchon in his seminal work, [8].

(d) If Φ is the reproducing kernel for a reproducing kernel Hilbert space $\mathcal{H}$

of functions $f : \Omega \to \mathbb{R}$, then Φ is positive semi-definite, since

$$\begin{aligned}
\sum_{\xi \in \Xi} \sum_{\eta \in \Xi} \alpha_\xi \alpha_\eta \Phi(\xi, \eta) &= \sum_{\xi \in \Xi} \sum_{\eta \in \Xi} \alpha_\xi \alpha_\eta \left\langle \Phi(\cdot, \xi), \Phi(\cdot, \eta) \right\rangle_{\mathcal{H}} \\
&= \left\langle \sum_{\xi \in \Xi} \alpha_\xi \Phi(\cdot, \xi), \sum_{\eta \in \Xi} \alpha_\eta \Phi(\cdot, \eta) \right\rangle_{\mathcal{H}} \\
&= \left\| \sum_{\xi \in \Xi} \alpha_\xi \Phi(\cdot, \xi) \right\|_{\mathcal{H}}^2 \geq 0.
\end{aligned}$$

Furthermore, Φ is positive definite if and only if the point evaluation functionals δ_x, $x \in \Omega$, are linearly independent in $\mathcal{H}^*$.

The types of kernels we are interested in in this thesis do not live in the Euclidean setting. Rather, our underlying space will be a certain type of Riemannian manifold.

2.2 Native Spaces

The situation in Example 2.1.2(d) can be reversed. A continuous positive definite kernel $\Phi : \Omega \times \Omega \to \mathbb{R}$ generates a reproducing kernel Hilbert space (RKHS) $\mathcal{N}_\Phi(\Omega)$, for which Φ is the reproducing kernel, called the *native space* for Φ. This is achieved by first setting $F_\Phi(\Omega) = \text{span}\,\{\Phi(\cdot, x) : x \in \Omega\}$ and equipping $F_\Phi(\Omega)$ with the inner product

$$\left\langle \sum_{j=1}^{M} \alpha_j \Phi(\cdot, x_j), \sum_{k=1}^{N} \beta_k \Phi(\cdot, y_k) \right\rangle_\Phi = \sum_{j=1}^{M} \sum_{k=1}^{N} \alpha_j \beta_k \Phi(x_j, y_j).$$

This gives a an inner product space (not necessarily complete, so often referred to as a "pre-Hilbert space") and the completion of $F_\Phi(\Omega)$ with respect to the norm $\| \cdot \|_\Phi$ is the native space for Φ, $\mathcal{N}_\Phi$. The details of this construction can be found in Schaback, [25], and Wendland, [29], which are adaptations of earlier work by Aronszajn, [1] and [2]. The main takeaways for our purposes are that the members of the native space are, despite being abstract elements of a closure, continuous functions, and that the kernel Φ is indeed the reproducing kernel for $\mathcal{N}_\Phi(\Omega)$. In conjunction with Example 2.1.2(d), this shows that there is a one-to-one correspondence between reproducing kernel Hilbert spaces and positive definite kernels.

We assume that Ω is a finite measure space, with measure μ. We also assume that there are real-valued continuous functions, $\{\varphi_\ell\}_{\ell\in\mathbb{N}}$, that form a complete orthonormal basis for $L_2(\Omega)$. Thus, every $f \in L_2(\Omega)$ has the expansion

$$f = \sum_{\ell=0}^{\infty} \widehat{f}_\ell \varphi_\ell,\ \ \widehat{f}_\ell = \langle f, \varphi_\ell\rangle_{L_2}.$$

(For example, spherical harmonics on the sphere $\mathbb{S}^2$ or the Wigner-D functions on the rotation group $SO(3)$; indeed, this will be the case whenever we have a compact Riemannian manifold.)

We consider symmetric kernels that have an expansion of the form

$$\Phi(x, y) = \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)\varphi_\ell(x)\varphi_\ell(y). \tag{2.2.1}$$

If the series in (2.2.1) converges uniformly, then Φ, being the uniform limit of continuous functions, is continuous. Conversely, if Φ is continuous then Mercer's Theorem guarantees that the series in (2.2.1) converges absolutely and uniformly (see [24, p.245]). In section 2.8, we'll give a sufficient condition on the coefficients of Φ to guarantee continuity of Φ.

If Φ is continuous and $\widehat{\phi}(\ell) > 0$ for all $\ell \in \mathbb{N}$, then Φ is positive definite, and the native space for Φ is

$$\mathcal{N}_\Phi(\Omega) = \left\{ f \in L_2(\Omega) :\ \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)^{-1} \left|\widehat{f}_\ell\right|^2 < \infty \right\}, \tag{2.2.2}$$

which is a Hilbert space with inner product

$$\langle f, g\rangle_{\mathcal{N}_\Phi(\Omega)} = \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)^{-1}\widehat{f}_\ell \widehat{g}_\ell.$$

We check directly that this is a RKHS with reproducing kernel Φ. If $g = \Phi(\cdot, x)$, then a straightforward calculation using the orthonormality of the functions φ_ℓ shows $\widehat{g}_\ell = \widehat{\phi}(\ell)\varphi_\ell(x)$. Hence, for $f \in \mathcal{N}_\Phi(\Omega)$,

$$\langle f, \Phi(\cdot, x)\rangle_{\mathcal{N}_\Phi(\Omega)} = \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)^{-1}\widehat{f}_\ell \widehat{\phi}(\ell)\varphi_\ell(x) = \sum_{\ell=0}^{\infty} \widehat{f}_\ell \varphi_\ell(x) = f(x).$$

(This proves it - it is a straightforward exercise to show that reproducing kernels are unique.)

If Φ is continuous and $\widehat{\phi}(\ell) > 0$ for all $\ell > L$, with $L \in \mathbb{N}$, then Φ is conditionally positive definite with respect to the auxiliary space $\Pi_L = \operatorname{span}\{\varphi_\ell : 0 \le \ell \le L\}$. The native space for Φ in this context is

$$\mathcal{N}_{\Phi,L}(\Omega) = \left\{ f \in L_2(\Omega) : \sum_{\ell=L+1}^{\infty} \widehat{\phi}(\ell)^{-1} \left|\widehat{f}_\ell\right|^2 < \infty \right\},$$

which is a semi-Hilbert space with semi-inner product

$$\langle f, g\rangle_{\mathcal{N}_{\Phi,L}(\Omega)} = \sum_{\ell=L+1}^{\infty} \widehat{\phi}(\ell)^{-1}\widehat{f}_\ell \widehat{g}_\ell.$$

We note that in contrast to the case of a positive definite kernel, we don't have reproduction; indeed, if $p \in \Pi_L$ then $\langle p, \Phi(\cdot, x)\rangle_{\mathcal{N}_{\Phi,L}(\Omega)} = 0$ for all $x \in \Omega$. If, however, Ξ is a finite Π_L-unisolvent subset of Ω and the coefficients $\{a_\xi\}_{\xi\in\Xi}$ annihilate $\Pi_L|_\Xi$, then we do have $\left\langle f, \sum_{\xi\in\Xi} a_\xi \Phi(\cdot, \xi)\right\rangle_{\mathcal{N}_{\Phi,L}(\Omega)} = \sum_{\xi\in\Xi} a_\xi f(\xi)$.

2.3 Sobolev Spaces I

We now insist on slightly more structure on our underlying space; namely, that it is a compact Riemannian manifold $\mathbb{M}$. In this setting the Laplace-Beltrami operator on $C^\infty(\mathbb{M})$, $-\Delta$, has countably many positive eigenvalues $\{\lambda_\ell\}_{\ell\in\mathbb{N}}$, corresponding to eigenfunctions $\{\varphi_\ell\}_{\ell\in\mathbb{N}}$ that form a complete orthonormal basis for $L_2(\mathbb{M})$. We are ignoring multiplicity here - the eigenvalues are not necessarily distinct, though the eigenfunctions are.

We now give our first definition of the term "Sobolev space" - $H^\tau(\mathbb{M})$. The advantage of this definition is that it doesn't require the computation of any covariant derivatives - members are determined solely based on the convergence of a series. (The covariant derivative will be introduced in section 2.5.) The drawback is that this definition can't be used to define Sobolev spaces on *subsets* of $\mathbb{M}$. We will explore a second definition in Section 2.6 - $W_2^m(\mathbb{M})$ - which provides a natural analog on subsets; subsequently, we will show that for nonnegative integers m there is a norm equivalence between the two seemingly different spaces $H^m(\mathbb{M})$ and $W_2^m(\mathbb{M})$. From here on, when the underlying space is the whole manifold, we will simply write H^m, W_2^m, C^∞, L_2, etc. instead of $H^m(\mathbb{M})$, $W_2^m(\mathbb{M})$, $C^\infty(\mathbb{M})$, $L_2(\mathbb{M})$, etc.

For $\tau \geq 0$ the *Sobolev space* H^τ is

$$H^\tau = \left\{ f \in L_2 : \sum_{\ell=0}^{\infty} (1+\lambda_\ell)^\tau \left| \widehat{f}_\ell \right|^2 \leq \infty \right\}, \tag{2.3.1}$$

which is a Hilbert space with inner product

$$\langle f, g \rangle_{H^\tau} = \sum_{\ell=0}^{\infty} (1+\lambda_\ell)^\tau \, \widehat{f}_\ell \widehat{g}_\ell$$

Given the similarities between the definitions (2.2.2) and (2.3.1), it is not hard to see that the particular kernel Φ of the form (2.2.1) with coefficients

$\widehat{\phi}(\ell)$ that are *exactly* $(1+\lambda_\ell)^{-\tau}$ is of interest. In the case that τ is a nonnegative integer, then Φ is the fundamental solution of $(1-\Delta)^\tau$. Indeed, there is a broader context here. If Q is a polynomial of degree τ and Φ is the kernel of the form (2.2.1) with coefficients $\widehat{\phi}(\ell) = Q(\lambda_\ell)^{-1}$, then Φ is the fundamental solution to $Q(\Delta)$. Formally this is justified by the following. Since Δ is self-adjoint and $\mathbb{M}$ has no boundary,

$$\begin{aligned}
\int_{\mathbb{M}} \left[Q(\Delta)f\right](y)\Phi(x,y)d\mu(y) &= \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)\varphi_\ell(x) \int_{\mathbb{M}} \left[Q(\Delta)f\right](y)\varphi_\ell(y)d\mu(y) \\
&= \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)\varphi_\ell(x) \int_{\mathbb{M}} f(y)\left[Q(\Delta)\varphi_\ell\right](y)d\mu(y) \\
&= \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)Q(\lambda_\ell)\varphi_\ell(x) \int_{\mathbb{M}} f(y)\varphi_\ell(y)d\mu(y) \\
&= \sum_{\ell=0}^{\infty} \widehat{f}_\ell\varphi_\ell(x) = f(x).
\end{aligned}$$

Remark 2.3.2. Switching the integral and the summation in the very first line above is purely formal. In practice this would have to be justified, say by the uniform convergence of (2.2.1), for example.

We end this subsection by noting that the eigenfunctions $\{\varphi_\ell\}_{\ell\in\mathbb{N}}$ are also orthogonal (though not necessarily orthonormal) with respect to the H^τ inner product - the proof is straightforward.

2.4 Kernel-Based Interpolation

We interpolate on a finite subset Ξ of $\mathbb{M}$. This means that we are given a set of values $c = \{c_\xi\}_{\xi\in\Xi} \in \mathbb{R}^\Xi$, and we want to find a function $s : \mathbb{M} \to \mathbb{R}$ such that $s(\xi) = c_\xi$ for each $\xi \in \Xi$. If Φ is positive definite, then the *approximation space* from which we obtain our interpolants is

$$V_{\Phi,\Xi} = \operatorname{span}\{\Phi(\cdot,\xi)\}_{\xi\in\Xi}.$$

To find the unique interpolant s from this space, we form the collocation matrix

$$K_\Xi = [\Phi(\xi,\eta)]_{\xi,\eta\in\Xi},$$

which is a symmetric positive definite matrix, and then find our interpolant

$$s = \sum_{\xi\in\Xi} a_\xi \Phi(\cdot,\xi)$$

by solving the linear system $K_\Xi a = c$ for the coefficients $a = \{a_\xi\}_{\xi\in\Xi}$.

Often the values come from a function $f \in \mathcal{N}_\Phi$; i.e. $c = \{f(\xi)\}_{\xi\in\Xi} = f|_\Xi$. In this case we view interpolation as an operator I_Ξ from $\mathcal{N}_\Phi$ to $V_{\Phi,\Xi}$; i.e., $s = I_\Xi f$. Indeed, $I_\Xi f$ is the orthogonal projection of $f \in \mathcal{N}_\Phi(\mathbb{M})$ onto $V_{\Phi,\Xi}$ in the native space inner product, since for all $\xi \in \Xi$,

$$\langle f - I_\Xi f, \Phi(\cdot,\xi)\rangle_{\mathcal{N}_\Phi} = f(\xi) - I_\Xi f(\xi) = 0.$$

This means we have a Pythagorean Theorem: $\|f\|^2_{\mathcal{N}_\Phi} = \|f - I_\Xi f\|^2_{\mathcal{N}_\Phi} + \|I_\Xi f\|^2_{\mathcal{N}_\Phi}$; consequently, we have the following.

Proposition 2.4.1. (Native Space Interpolation Error Estimate for Positive Definite Kernels) *Suppose Φ is positive definite and Ξ is a finite subset of $\mathbb{M}$. Let $f \in \mathcal{N}_\Phi(\mathbb{M})$ and let $I_\Xi f$ be the interpolant to f in V_Ξ. Then*

$$\|f - I_\Xi f\|_{\mathcal{N}_\Phi(\mathbb{M})} \le \|f\|_{\mathcal{N}_\Phi(\mathbb{M})}.$$

Now suppose Φ is conditionally positive definite with respect to Π_L, and that Ξ is a finite, Π_L-unisolvent subset of $\mathbb{M}$. The matrix K_Ξ is no longer guaranteed to be positive definite, or indeed even invertible. Let $Q = \dim(\Pi_L)$ and $\{p_1, \ldots, p_Q\}$ be a basis for Π_L, and construct the matrix $P = [p_k(\xi)]_{\xi\in\Xi, k\in\{1,\ldots,Q\}}$. Then the augmented collocation matrix

$$K_{\Xi,L} = \begin{bmatrix} K_\Xi & P \\ P^T & 0 \end{bmatrix}, \tag{2.4.2}$$

is positive definite. For a conditionally positive definite kernel, the interpolant is

$$s = \sum_{\xi\in\Xi} a_\xi \Phi(\cdot, \xi) + \sum_{k=1}^{Q} b_k p_k,$$

where the coefficients $a = \{a_\xi\}_{\xi\in\Xi}$ and $b = \{b_k\}_{k\in\{1,\ldots,Q\}}$ are obtained by solving the linear system

$$K_{\Xi,L} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} c \\ 0 \end{bmatrix}.$$

The *approximation space* in the conditionally positive definite setting is

$$V_{\Phi,\Xi,L} = \left\{ \sum_{\xi\in\Xi} a_\xi \Phi(\cdot, \xi) : \sum_{\xi\in\Xi} a_\xi p(\xi) = 0 \ \forall\, p \in \Pi_L \right\} + \Pi_L.$$

We say in this context that for $\sum_{\xi\in\Xi} a_\xi \Phi(\cdot, \xi) + p \in V_{\Phi,\Xi,L}$, the coefficients $a = \{a_\xi\}_{\xi\in\Xi}$ *annihilate* $\Pi_L|_\Xi$.

As before, if the values come from a function $f \in \mathcal{N}_{\Phi,L}$, then we view interpolation as an operator $I_{\Xi,L}$ from $\mathcal{N}_{\Phi,L}$ to $V_{\Xi,L}$; i.e., $s = I_{\Xi,L} f$. We note that in this situation interpolation reproduces the auxiliary space Π_L; i.e., $I_{\Xi,L} p = p$ for all $p \in \Pi_L$. It is also important to note that just as before, $I_{\Xi,L} f$ is the orthogonal projection of $f \in \mathcal{N}_{\Phi,L}(\mathbb{M})$ to $V_{\Xi,L}$ in the semi-inner product $\langle \cdot, \cdot \rangle_{\mathcal{N}_{\Phi,L}(\mathbb{M})}$. This again means that we have a "Pythagorean Theorem" $\|f\|^2_{\mathcal{N}_{\Phi,L}} = \|f - I_{\Xi,L} f\|^2_{\mathcal{N}_{\Phi,L}} + \|I_{\Xi,L} f\|^2_{\mathcal{N}_\Phi}$, and again we have the following consequence.

Proposition 2.4.3. (Native Space Interpolation Error Estimate for Conditionally Positive Definite Kernels) *Suppose Φ is conditionally positive definite with respect to Π_L and Ξ is a finite Π_L-unisolvent subset of $\mathbb{M}$. Let $f \in \mathcal{N}_{\Phi,L}(\mathbb{M})$ and let $I_{\Xi,L} f$ be the interpolant to f in $V_{\Xi,L}$. Then*

$$\|f - I_{\Xi,L} f\|_{\mathcal{N}_{\Phi,L}(\mathbb{M})} \leq \|f\|_{\mathcal{N}_{\Phi,L}(\mathbb{M})}.$$

The fact that the interpolation operator is the orthogonal projector of $\mathcal{N}_\Phi$ onto V_Ξ when Φ is positive definite has another consequence. Since $I_\Xi f$ is the *unique* element u of $V_{\Phi,\Xi}$ such that $u|_\Xi = f|_\Xi$, if g is any other element of $\mathcal{N}_\Phi$ for which $g|_\Xi = f|_\Xi$, then $I_\Xi g = I_\Xi f$. Thus, the "Pythagorean Theorem" $\|g\|^2_{\mathcal{N}_\Phi} = \|g - I_\Xi g\|^2_{\mathcal{N}_\Phi} + \|I_\Xi g\|^2_{\mathcal{N}_\Phi}$ also gives that

$$\|I_\Xi f\|_{\mathcal{N}_\Phi} = \|I_\Xi g\|_{\mathcal{N}_\Phi} \leq \|g\|_{\mathcal{N}_\Phi}.$$

We summarize in the following result.

Proposition 2.4.4. *If Φ is positive definite and $f \in \mathcal{N}_\Phi$, then for every finite subset Ξ of $\mathbb{M}$, $I_\Xi f$ is the unique native space norm-minimizing member u of $\mathcal{N}_\Phi$ such that $u|_\Xi = f|_\Xi$.*

Similar considerations gives an analogous result for $I_{\Xi,L}$.

Proposition 2.4.5. *If Φ is conditionally positive definite with respect to Π_L and $f \in \mathcal{N}_{\Phi,L}$, then for every finite Π_L-unisolvent subset of $\mathbb{M}$, $I_{\Xi,L} f$ is the unique native space semi-norm-minimizing member u of $\mathcal{N}_{\Phi,L}$ such that $u|_\Xi = f|_\Xi$.*

2.5 Calculus on Manifolds

The geometric background we need is developed in [6] and [7], and detailed in [15] and [16]. We give a brief summary here. We assume $\mathbb{M}$ is a compact C^∞ Riemannian manifold without boundary. Let g be the Riemannian metric of $\mathbb{M}$.

The metric g defines an inner product $\langle \cdot, \cdot \rangle_{g,p}$ on each tangent space $T_p\mathbb{M}$. We will usually omit the metric g and simply write $\langle \cdot, \cdot \rangle_p$ to denote the metric at $p \in \mathbb{M}$. An order k *covariant tensor* is a real-valued multilinear function of the k-fold product of $T_p\mathbb{M}$. Thus, the set of order k covariant tensors is the k-fold tensor product $\left(T_p^*\mathbb{M}\right)^{\otimes k} = T_p^*\mathbb{M} \otimes \cdots \otimes T_p^*\mathbb{M}$. In coordinates, there is a smoothly varying basis $\left\{\mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k}\right\}_{\hat{\imath} \in \{1,\ldots,d\}^k}$, where we adopt the convention $\hat{\imath} = (i_1, \ldots, i_k)$. So, a covariant tensor $\mathbf{T}$ can be written as

$$\mathbf{T} = \sum_{\hat{\imath} \in \{1,\ldots,d\}^k} T_{\hat{\imath}} \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k},$$

and the $T_{\hat{\imath}}$ are called the *covariant components* of $\mathbf{T}$. The metric g is itself an order 2 covariant tensor. Similarly, an order k contravariant tensor is an element of the k-fold tensor product $(T_p\mathbb{M})^{\otimes k} = T_p\mathbb{M} \otimes \ldots \otimes T_p\mathbb{M}$. One can also define tensors of mixed type.

The *covariant derivative*, or *connection*, ∇, is defined as follows. If $\mathbf{T}$ is an order k covariant tensor as above, then

$$\nabla \mathbf{T} = \sum_{\hat{\imath} \in \{1,\ldots,d\}^k} \sum_{j=1}^{d} (\nabla \mathbf{T})_{\hat{\imath},j}\, \mathbf{e}^{i_1} \otimes \cdots \otimes \mathbf{e}^{i_k} \otimes \mathbf{e}^j,$$

where the covariant components of $\nabla\mathbf{T}$ are given by

$$(\nabla\mathbf{T})_{\hat{\imath},j} = \frac{\partial T_{\hat{\imath}}}{\partial x^j} - \sum_{r=1}^{k}\sum_{s=1}^{d}\Gamma^s_{j,i_r} T_{i_1,\ldots,i_{r-1},s,i_{r+1},\ldots,i_k}.$$

Here, the $\Gamma^k_{i,j}$'s are Christoffel symbols. Thus, if $\mathbf{T}$ is an order k tensor, $\nabla\mathbf{T}$ is an order $k+1$ tensor.

The metric g induces an invariant inner product on $\left(T^*_p\mathbb{M}\right)^{\otimes k}$:

$$\langle \mathbf{T}, \mathbf{S}\rangle_p = \sum_{\hat{\imath},\hat{\jmath}\in\{1,\ldots,d\}^k} g^{i_1 j_1}\cdots g^{i_k j_k} S_{\hat{\imath}} T_{\hat{\jmath}}. \tag{2.5.1}$$

The *adjoint* ∇^* of ∇ is defined by

$$\int_{\mathbb{M}} \langle \nabla\mathbf{T}, \mathbf{S}\rangle_p \, d\mu(p) = \int_{\mathbb{M}} \langle \mathbf{T}, \nabla^*\mathbf{S}\rangle_p \, d\mu(p),$$

whenever T is a rank k tensor and S is a rank $k+1$ tensor. So, it takes order $k+1$ tensors to order k tensors.

A smooth function $f : \mathbb{M} \longrightarrow \mathbb{R}$ is an order 0 tensor; whereas ∇f is an order 1 (covariant) tensor. In general, $\nabla^k f$ is an order k tensor. In coordinates $(\mathfrak{U}, \phi)$, where $\phi : \mathfrak{U} \to U \subset \mathbb{R}^d$ is a chart and $\phi(p) = \left(x^1, \ldots, x^d\right) \in U$, the components of $\nabla^k f$ are

$$\left(\nabla^k f(x)\right)_{\hat{\imath}} = \left(\partial^k f(x)\right)_{\hat{\imath}} + \sum_{m=1}^{k-1}\sum_{\hat{\jmath}\in\{1,\ldots,d\}^m} A^{\hat{\jmath}}_{\hat{\imath}}(x)\left(\partial^m f(x)\right)_{\hat{\jmath}},$$

where

$$\left(\partial^m f(x)\right)_{\hat{\jmath}} = \frac{\partial^m}{\partial x^{j_1}\cdot\partial x^{j_m}} f\circ\phi^{-1}.$$

Here the coefficients $A^{\hat{\jmath}}_{\hat{\imath}}(x)$ depend on the Christoffel symbols and their derivatives to order $k-1$, and hence are smooth in x.

Applying ∇^* to the order k tensor $\nabla^k f$ k times produces a scalar,

$(\nabla^*)^k \nabla^k f$, which is the same as $\left(\nabla^k\right)^* \nabla^k f$. When $k = 1$, we have the *Laplace-Beltrami* operator, $-\Delta = \nabla^* \nabla$. In coordinates,

$$\Delta f = \frac{1}{\sqrt{\det(g)}} \sum_{j=1}^{d} \sum_{k=1}^{d} \frac{\partial}{\partial x^j} \left[g^{jk} \sqrt{\det(g)} \frac{\partial f}{\partial x^k} \right].$$

Crucially, it is *not* necessarily the case that for $k > 1$, Δ^k is equal to $(-1)^k \left(\nabla^k\right)^* \nabla^k$, though it is the case if the manifold is flat (like Euclidean space or the flat torus). This is because ∇ and ∇^* do not necessarily commute in general. For example, $\Delta^2 = \nabla^* \nabla \nabla^* \nabla$, whereas $\left(\nabla^2\right)^* \nabla^2 = \left(\nabla^*\right)^2 \nabla^2 = \nabla^* \nabla^* \nabla \nabla$, and these are not necessarily the same.

We do, however, have the following propositions, which will allow us to prove a norm equivalence between our different notions of Sobolev spaces in the next section. The first is due to Helgason - [18] - but is stated here as it appears in [16, Lemma 4.2 and Assumption 4.1]. We state both and prove the second in the special case that $\mathbb{M}$ is *two-point homogeneous*, but it holds for all compact manifolds without boundary. Two point homogeneity means that for every two pairs of points p, q and p', q' in $\mathbb{M}$ with $\mathrm{dist}(p, q) = \mathrm{dist}(p', q')$, there is an isometry $\psi : \mathbb{M} \to \mathbb{M}$ with $\psi(p) = p'$ and $\psi(q) = q'$. Intuitively, this means that the manifold $\mathbb{M}$ "looks the same" at every point.

Proposition 2.5.2. *Let $\mathbb{M}$ be a two-point homogeneous space. Then for all $k \in \mathbb{N}$ there is a real polynomial p_{k-1} of degree $k - 1$ such that*

$$\left(\nabla^k\right)^* \nabla^k = (-1)^k \Delta^k + p_{k-1}(\Delta).$$

Proposition 2.5.3. *Let $\mathbb{M}$ be a two-point homogeneous space. If $Q(x) = c_m x^m + \ldots + c_0$ is a real polynomial of degree m, then there exist real* *numbers* a_j, $j \in \{0, \ldots, m\}$ *such that*

$$Q(\Delta) = \sum_{j=0}^{m} a_j \left(\nabla^j\right)^* \nabla^j.$$

Proof. Clearly it suffices to show the proposition is true when $Q(x) = x^k$ for $k \in \mathbb{N}$. We proceed by strong induction. The base case $m = 1$ is trivial, since $\Delta = -\nabla^*\nabla$. Suppose the proposition is true for $k \in \{1, \ldots, m\}$. By Proposition 2.5.2,

$$\left(\nabla^{m+1}\right)^* \nabla^{m+1} = (-1)^{m+1}\Delta^{m+1} + p_m(\Delta)$$

for a polynomial p_m of degree m. Let $p_m(x) = \sum_{k=0}^{m} b_k x^k$. By our strong inductive hypothesis, for each $k \in \{0, \ldots, m\}$ there are numbers $c_{k,j}$ such that $\Delta^k = \sum_{j=0}^{k} c_{k,j} \left(\nabla^k\right)^* \nabla^k$. Hence,

$$\begin{aligned}
\Delta^{m+1} &= (-1)^{m+1} \left(\nabla^{m+1}\right)^* \nabla^{m+1} - (-1)^{m+1} \sum_{k=0}^{m} b_k \sum_{j=0}^{k} c_{k,j} \left(\nabla^j\right)^* \nabla^j \\
&= (-1)^{m+1} \left(\nabla^{m+1}\right)^* \nabla^{m+1} + \sum_{j=0}^{m} \left((-1)^m \sum_{k=j}^{m} b_k c_{k,j} \right) \left(\nabla^j\right)^* \nabla^j.
\end{aligned}$$

Hence we can take $a_{m+1} = (-1)^{m+1}$ and $a_j = (-1)^m \sum_{k=j}^{m} b_k c_{k,j}$ for $j \in \{0, \ldots, m\}$. ☐

2.6 Sobolev Spaces II

Now that we have the covariant derivative, we are ready to give our second definition of the term "Sobolev space". For an integer m and a measurable subset Ω of $\mathbb{M}$, the *Sobolev space* $W_2^m(\Omega)$ is

$$W_2^m(\Omega) = \left\{ f \in L_2(\Omega) : \sum_{k=0}^{m} \int_{\Omega} \left\langle \nabla^k f, \nabla^k f \right\rangle_p d\mu(p) < \infty \right\},$$

where the inner product in the integrand is given in (2.5.1). It is a Hilbert space with inner product

$$\langle f, g \rangle_{W_2^m(\Omega)} = \sum_{k=0}^{m} \int_{\Omega} \left\langle \nabla^k f, \nabla^k g \right\rangle_p d\mu(p).$$

Again, when Ω is all of $\mathbb{M}$, we denote it simply by W_2^m instead of $W_2^m(\mathbb{M})$.

An important tool for obtaining interpolation error estimates is the Zeros Lemma, which will appear in the next section. We'd like to have a similar result on our Sobolev spaces, and [15, Lemma 3.2] allows us to do so.

Before we state [15, Lemma 3.2], we need to define the term "injectivity radius". At a point p on a Riemannian manifold $\mathbb{M}$, the *exponential map* at p, $\mathrm{Exp}_p : T_x\mathbb{M} \to \mathbb{M}$, is defined as follows. For each $v \in T_p\mathbb{M}$ there is a unique geodesic γ_v in $\mathbb{M}$ such that $\gamma(0) = p$ and $\gamma'(0) = v$. Define $\exp_p(v) = \gamma_v(1)$. The exponential map is in general not a global diffeomorphism; however, at the origin of the tangent space, the differential of Exp_p is the identity. So, the Inverse Function Theorem guarantees that there is a neighborhood of the origin on which Exp_p is injective. The supremum of all radii r for which Exp_p is injective on $B(0, r) \subset T_p(\mathbb{M})$ is called the *injectivity radius* of $\mathbb{M}$ at p, denoted r_p. The *injectivity radius* of $\mathbb{M}$, denoted $r_{\mathbb{M}}$, is the infimum of r_p as p ranges over all of $\mathbb{M}$. If $0 < r_{\mathbb{M}} \leq \infty$, then $\mathbb{M}$ is said to have positive injectivity radius. This is the case for compact manifolds.

Lemma 2.6.1. ([15, Lemma 3.2]) *Let $m \in \mathbb{N}$ and $0 < r < r_{\mathbb{M}}/3$. There are constants $0 < c_1 \leq c_2$ so that for any measurable $\Omega \subset B(0, r)$, all $j \in \mathbb{N}$*

with $j \leq m$, and any $p \in \mathbb{M}$, the equivalence

$$c_1 \left\| u \circ \operatorname{Exp}_p \right\|_{W_2^j(\Omega)} \leq \|u\|_{W_2^j\left(\operatorname{Exp}_p(\Omega)\right)} \leq c_2 \left\| u \circ \operatorname{Exp}_p \right\|_{W_2^j(\Omega)}$$

holds for all $u : \operatorname{Exp}_x(\Omega) \to \mathbb{R}$ in $W_2^m\left(\operatorname{Exp}_p(\Omega)\right)$. The constants c_1 and c_2 depend on r and m but are independent of Ω and p.

For an open subset Ω of $\mathbb{R}^d$ that is bounded and has Lipschitz boundary, the Sobolev embedding theorem guarantees that $W_2^m(\Omega) \subset C(\Omega)$ for a subset Ω of $\mathbb{R}^d$ when $m > d/2$. Lemma 2.6.1 says that this holds on manifolds as well; if $m > d/2$ and $\Omega \subset \mathbb{M}$ is open, bounded, and has Lipschitz boundary, then $W_2^m(\Omega) \subset C(\Omega)$.

2.7 Sets of Centers; The Zeros Lemma

At this point we need to introduce three quantities that measure the so-called "uniformity" of our set of centers Ξ (sometimes called nodes). The first is the *fill distance* (sometimes called the *mesh norm*),

$$h_\Xi = \max_{p \in \mathbb{M}} \operatorname{dist}(p, \Xi) = \max_{p \in \mathbb{M}} \min_{\xi \in \Xi} \operatorname{dist}(p, \xi).$$

It measures the density of the centers in $\mathbb{M}$. The second is the *separation distance*,

$$q_\Xi = \frac{1}{2} \min_{\xi \in \Xi} \operatorname{dist}(\xi, \Xi \setminus \xi) = \frac{1}{2} \min_{\xi \in \Xi} \min_{\substack{\eta \in \Xi \\ \eta \neq \xi}} \operatorname{dist}(\xi, \eta).$$

It measures how evenly distributed the centers are. The third is the *mesh ratio*, $\rho_\Xi = h_\Xi / q_\Xi$. Note that the mesh ratio is at least 2, since the definitions force $h_\Xi \geq 2q_\Xi$. Sets of centers for which the mesh ratio is bounded by some fixed ρ (preferably close to 2) are sometimes called *quasi-uniform*. We will assume throughout the remainder of this paper that our sets of centers are quasi-uniform with mesh ratios controlled above by a fixed $\rho \geq 2$. In the sequel we will allow the constants in our estimates to depend on ρ. We note that in the context of a finite, quasi-uniform set of centers Ξ in a compact d-dimensional manifold $\mathbb{M}$, there are constants C_1 and C_2 such that $C_1 h_\Xi^{-d} \leq \#\Xi \leq C_2 q_\Xi^{-d}$. We also note that $h_\Xi = \rho q_\Xi$, and therefore, since we allow the constants in our estimates to depend on ρ, h_Ξ and q_Ξ are interchangeable.

We now state the Zeros Lemma, which, roughly speaking, says that a weak norm of a function with many zeros can be controlled by a stronger norm with a multiplicative constant that decreases with the density of the zeros. A zeros lemma was first stated and proved in [8], where the underlying space was a Euclidean ball. Narcowich, Ward, and Wendland extend the result to a broader class of spaces in [23]. In [16, Appendix A], Hangelbroek, Narcowich, and Ward extend it to manifolds, which is the version we use.

Lemma 2.7.1. (The Zeros Lemma for Manifolds - [16, Corollary A.13]) *Let $m, k \in \mathbb{N}$ satisfy $m > d/2$, $0 \leq k \leq m$. In addition, let Ξ be a finite subset of $\mathbb{M}$. If $u \in W_2^m$ satisfies $u|_\Xi = 0$, then for h_Ξ sufficiently small,*

$$\|u\|_{W_2^k} \leq C h_\Xi^{m-k} \|u\|_{W_2^m},$$

and

$$\|u\|_{L_\infty} \leq C h_\Xi^{m-\frac{d}{2}} \|u\|_{W_2^m}.$$

with a constant C that is independent of Ξ and u.

We now show the norm equivalence between H^m and W_2^m when m is a nonnegative integer. Two important points to note are that

$$\langle f, g\rangle_{H^m} = \langle (I-\Delta)^m f, f\rangle_{L_2},$$

whereas

$$\langle f, g\rangle_{W_2^m} = \sum_{k=0}^{m} \left\langle \nabla^k f, \nabla^k f\right\rangle_{L_2}.$$

Remark 2.7.2. We include the following result, and its proof, in the interest of self-containment. It is a simplified version of a much broader result found in [28]. Let us add a bit of historical context. Aubin introduced the W_2^m Sobolev spaces on Riemannian manifolds in [3], whereas Strichartz introduced the H^m Sobolev spaces on Riemannian manifolds in [27]. In [28], Triebel showed these spaces coincide for much larger and more general classes of function spaces than those that we will be considering. The full result found in [28] is, of course, true, but requires a far greater amount of analysis than is necessary for our context.

Proposition 2.7.3. [28, Theorem 4.1] *Suppose* $\mathbb{M}$ *is a compact Riemannian manifold without boundary. For a nonnegative integer* m*, the spaces* $H^m(\mathbb{M})$ *and* $W_2^m(\mathbb{M})$ *are the same, and the norms* $\|\cdot\|_{H^m}$ *and* $\|\cdot\|_{W_2^m}$ *are equivalent.*

Proof in the case that $\mathbb{M}$ *is two-point homogeneous.* Proposition 2.5.3 with $Q(x) = (1-x)^m$ gives real numbers a_j, $j = 0, \ldots, m$ for which $(1-\Delta)^m = \sum_{j=0}^{m} a_j \left(\nabla^j\right)^* \nabla^j$. This gives

$$\begin{aligned}\|f\|_{H^m}^2 &= \sum_{j=0}^{m} a_j \left\langle \left(\nabla^j\right)^* \nabla^j f, f\right\rangle_{L_2} \\ &\le C \sum_{j=0}^{m} \left\langle \nabla^j f, \nabla^j f\right\rangle_{L_2} = C\|f\|_{W_2^m},\end{aligned}$$

where we have taken $C = \max_{j\in\{0,\ldots,m\}} |a_j|$.

For the other direction, we start by using Theorem 2.5.2 to obtain

$$\|f\|_{W_2^m}^2 = \sum_{k=0}^{m} \left\langle \left((-1)^k \Delta^k + p_{k-1}(\Delta)\right) f, f\right\rangle_{L_2}. \tag{2.7.4}$$

Now, $\sum_{k=0}^{m}\left((-1)^k x^k + p_{k-1}(x)\right)$ is a polynomial of degree m. Write it as $\sum_{j=0}^{m} b_j x^j$ to obtain

$$\|f\|_{W_2^m} = \sum_{j=0}^{m} b_j \left\langle \Delta^j f, f\right\rangle_{L_2} \leq \sum_{j=0}^{m} |b_j| \left|\left\langle \Delta^j f, f\right\rangle_{L_2}\right|.$$

Note that since $-\Delta$ is positive definite, $\left|\left\langle \Delta^j f, f\right\rangle_{L_2}\right| = \left\langle (-1)^j \Delta^j f, f\right\rangle_{L_2}$ for each j. Letting $C = \max_{j\in\{0,\ldots,m\}} |b_j|$, we arrive at

$$\|f\|_{W_2^m} \leq C \sum_{j=0}^{m} \left\langle (-1)^j \Delta^j f, f\right\rangle_{L_2}.$$

But $\left\langle (-1)^j \Delta^j f, f\right\rangle_{L_2} \leq \left\langle (I-\Delta)^j f, f\right\rangle_{L_2}$, since the term on the left is but one of the positive terms that appears in the binomial expansion of the term on the right. Hence,

$$\|f\|_{W_2^m} \leq C \sum_{j=0}^{m} \left\langle (I-\Delta)^j f, f\right\rangle_{L_2} = C \sum_{j=0}^{m} \|f\|_{H_j} \leq C(m+1)\|f\|_{H^m},$$

where in the last inequality we have used the fact that $\|\cdot\|_{H_j} \leq \|\cdot\|_{H^m}$ for $0 \leq j \leq m$. □

We end this subsection with a Lemma from [5] that deals with the Sobolev norms of products of functions in Sobolev spaces which will be useful in the sequel. The version of this which is used here is from Couldon, et al. - it extends to manifolds an earlier result (the "generalized Leibniz rule") on $\mathbb{R}^d$ by Gulisashvili and Kon, [13, Theorem 1.4].

Lemma 2.7.5. ([5, Theorem 27]) *Let f, g be in $H^m \cap L_\infty$, where $m \in \mathbb{N}$ satisfies $m > d/2$. Then $fg \in H^m \cap L_\infty$ and there exists a $C > 0$ such that*

$$\|fg\|_{H^m} \leq C\left(\|f\|_{H^m}\|g\|_{L_\infty} + \|f\|_{L_\infty}\|g\|_{H^m}\right).$$

2.8 Interpolation Error Estimates

We proceed by adding a further assumption, this time not on the manifold $\mathbb{M}$, but on the coefficients of the kernel Φ. Specifically, we assume that Φ is a kernel of the form (2.2.1) whose coefficients satisfy

$$C_1 (1 + \lambda_\ell)^{-\tau} \leq \widehat{\phi}(\ell) \leq C_2 (1 + \lambda_\ell)^{-\tau}. \tag{2.8.1}$$

for $\tau > d/2$, either for all $\ell \in \mathbb{N}$ if Φ is positive definite, or for all $\ell > L \in \mathbb{N}$ if Φ is conditionally positive definite with respect to Π_L.

In the case where Φ is positive definite, it is clear from the definitions (2.3.1) and (2.2.2) that the native space norm $\|\cdot\|_{\mathcal{N}_\Phi}$ and the Sobolev norm $\|\cdot\|_{H^\tau}$ are equivalent. This, along with the Zeros Lemma, is all we need to prove our interpolation error estimate in the case of a positive definite kernel. We note first that if $\tau = m \in \mathbb{N}$ with $m > d/2$, then

$$\mathcal{N}_\Phi = H^m = W_2^m \subset C(\mathbb{M}).$$

(The last equality is due to the Sobolev embedding theorem.)

Theorem 2.8.2. *Suppose Φ is a positive definite kernel of the form (2.2.1) whose coefficients satisfy (2.8.1) for some $\tau \in \mathbb{N}$ with $\tau > d/2$ and all $\ell \in \mathbb{N}$, and let $\sigma \in \mathbb{N}$ with $0 \leq \sigma \leq \tau$. Let Ξ be a finite subset of $\mathbb{M}$, $f \in H^\tau$, and $I_\Xi f$ the interpolant to f in $V_{\Phi,\Xi}$. There exists a constant C which is independent of f and Ξ such that, for h_Ξ sufficiently small,*

$$\|f - I_\Xi f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f\|_{H^\tau}.$$

Proof. By the Zeros Lemma,

$$\|f - I_\Xi f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f - I_\Xi f\|_{H^\tau}.$$

Since the coefficients of Φ satisfy (2.8.1), the norms $\|\cdot\|_{H^\tau}$ and $\|\cdot\|_{\mathcal{N}_\Phi}$ are equivalent, and so

$$\|f - I_\Xi f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f - I_\Xi f\|_{\mathcal{N}_\Phi}.$$

By Proposition 2.4.1, then,

$$\|f - I_\Xi f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f\|_{\mathcal{N}_\Phi}.$$

Finally, using again the equivalence of those norms, we have

$$\|f - I_\Xi f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f\|_{H^\tau}.$$

□

For a kernel that is conditionally positive definite with respect to Π_L, and whose coefficients satisfy (2.8.1) for some τ and all $\ell > L$, the proof of the error estimate is a bit more involved, though the result itself is the same. The reason for this is that the native space semi-norm isn't equivalent to the Sobolev norm. There is a remedy, however, involving orthogonal projections.

We require the orthogonal projection P_L which projects all the Sobolev spaces onto Π_L. The projection is the following: if $f = \sum_{\ell=0}^{\infty} \widehat{f}_\ell \varphi_\ell \in L_2$, then $P_L f = \sum_{\ell=0}^{L} \widehat{f}_\ell \varphi_\ell$. This operator is the orthogonal projection onto Π_L, because if $f \in L_2$ and $\ell \in \{0, \ldots, L\}$, then $\widehat{(P_L f)}_\ell = \widehat{f}_\ell$, and if $g \in \Pi_L$, then $\widehat{g}_\ell = 0$ for $\ell > L$. Therefore if $f \in H_\sigma$ for any $\sigma \geq 0$,

$$\langle f - P_L f, g \rangle_{H_\sigma} = \sum_{\ell=0}^{L} (1+\lambda_\ell)^\sigma \left(\widehat{f}_\ell - \widehat{(P_L f)}_\ell \right) \widehat{g}_\ell = 0.$$

Our interpolation error estimate will rely on two facts. First, for $f \in H^\tau$, we have by (2.8.1) that

$$\|(I - P_L) f\|_{H^\tau}^2 = \sum_{\ell=L+1}^{\infty} (1+\lambda_\ell)^\tau \left| \widehat{f}_\ell \right|^2 \leq C_1^{-1} \sum_{\ell=L+1}^{\infty} \widehat{\phi}(\ell) \left| \widehat{f}_\ell \right|^2 = C \|f\|_{\mathcal{N}_{\Phi,L}}^2 . \tag{2.8.3}$$

Second, if $0 \leq \sigma \leq \tau$ and $f \in H^\tau$, then

$$\|P_L f\|_{H^\tau}^2 = \sum_{\ell=0}^{L} (1+\lambda_\ell)^\tau \left| \widehat{f}_\ell \right|^2 \leq C_{\sigma,\tau,L} \sum_{\ell=0}^{L} (1+\lambda_\ell)^\sigma \left| \widehat{f}_\ell \right|^2 = C_{\sigma,\tau,L} \|P_L f\|_{H_\sigma}^2 , \tag{2.8.4}$$

where $C_{\sigma,\tau,L} = (1+\lambda_L)^{\tau-\sigma}$ depends only on σ, τ, and L.

Theorem 2.8.5. *Suppose Φ is a kernel of the form (2.2.1) whose coefficients satisfy (2.8.1) for $\tau \in \mathbb{N}$ with $\tau > d/2$ and all $\ell > L \in \mathbb{N}$, and is conditionally positive definite with respect to Π_L, and let $\sigma \in \mathbb{N}$ with $0 \leq \sigma \leq \tau$. Let Ξ be a finite Π_L-unisolvent subset of $\mathbb{M}$, $f \in H^\tau$, and $I_{\Xi,L} f$ the interpolant to f in $V_{\Xi,L}$. There exists a constant C which is independent of f and Ξ such that, for h_Ξ sufficiently small,*

$$\|f - I_{\Xi,L} f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f\|_{H^\tau} .$$

Proof. The case $\sigma = \tau$ is just Proposition 2.4.3, so assume $\sigma < \tau$. By the Zeros Lemma,

$$\|f - I_{\Xi,L} f\|_{H_\sigma} \leq C h_\Xi^{\tau-\sigma} \|f - I_{\Xi,L} f\|_{H^\tau} .$$

By "smuggling in" both $P_L f$ and $P_L I_{\Xi,L} f$, we obtain

$$\begin{aligned}\|f - I_{\Xi,L} f\|_{H_\sigma} \le C h_\Xi^{\tau-\sigma} \big(&\|P_L (f - I_{\Xi,L} f)\|_{H^\tau} \\ &+ \|(I - P_L)(f - I_{\Xi,L} f)\|_{H^\tau}\big).\end{aligned}$$

Now,

$$\|P_L (f - I_{\Xi,L} f)\|_{H^\tau} \le C \|P_L (f - I_{\Xi,L} f)\|_{H_\sigma}$$

by (2.8.4), and

$$\|(I - P_L)(f - I_{\Xi,L} f)\|_{H^\tau} \le \|f - I_{\Xi,L} f\|_{\mathcal{N}_{\Phi,L}}$$

by (2.8.3). Hence,

$$\|f - I_{\Xi,L} f\|_{H_\sigma} \le C h_\Xi^{\tau-\sigma} \|f - I_{\Xi,L} f\|_{H_\sigma} + C h_\Xi^{\tau-\sigma} \|f - I_{\Xi,L} f\|_{\mathcal{N}_{\Phi,L}}.$$

By Proposition 2.4.3, $\|f - I_{\Xi,L} f\|_{\mathcal{N}_{\Phi,L}} \le \|f\|_{\mathcal{N}_{\Phi,L}}$. We can assume h_Ξ is small enough that $C h_\Xi^{\tau-\sigma} \le 1/2$. Thus,

$$\|f - I_{\Xi,L} f\|_{H_\sigma} \le \frac{1}{2} \|f - I_{\Xi,L} f\|_{H_\sigma} + C h_\Xi^{\tau-\sigma} \|f\|_{\mathcal{N}_{\Phi,L}}.$$

Subtracting $\frac{1}{2} \|f - I_{\Xi,L} f\|_{H_\sigma}$ from both sides and multiplying by 2 gives

$$\begin{aligned}\|f - I_{\Xi,L} f\|_{H_\sigma} &\le C h_\Xi^{\tau-\sigma} \|f\|_{\mathcal{N}_{\Phi,L}} = C h_\Xi^{\tau-\sigma} \sum_{\ell=L+1}^{\infty} \widehat{\phi}(\ell)^{-1} \left|\widehat{f}_\ell\right|^2 \\ &\le C C_2 h_\Xi^{\tau-\sigma} \sum_{\ell=L+1}^{\infty} (1+\lambda_\ell)^\tau \left|\widehat{f}_\ell\right|^2 \\ &\le C h_\Xi^{\tau-\sigma} \sum_{\ell=0}^{\infty} (1+\lambda_\ell)^\tau \left|\widehat{f}_\ell\right|^2 = C h_\Xi^{\tau-\sigma} \|f\|_{H^\tau}.\end{aligned}$$

(Here C_2 is the constant from (2.8.1).) □

Chapter 3

Kernel-Based Galerkin Methods

In this chapter we flush out the details for our kernel-based Galerkin methods. In section 3.1, we exhibit a quadrature rule that will be used to approximate integrals that appear as entries of the stiffness matrix at the heart of our Galerkin method. It is a weighted sum of samples of the integrand, analogous to Simpson's rule or the trapezoid rule in Calculus. The error estimate for our quadrature formula is a direct consequence of our interpolation error estimates. In section 3.2, we introduce the Lagrange basis, which is the basis for our approximation space in which we express our Galerkin approximations. In section 3.3 we give the details for the Galerkin method. Finally, we end the chapter by proving salient features of the stiffness matrix in section 3.4.

We maintain all of our assumptions, both on the kernel Φ and the manifold $\mathbb{M}$. In particular, that $\mathbb{M}$ is a compact C^∞ Riemannian manifold with no boundary, and that our kernel Φ is of the form (2.2.1) with coefficients that satisfy (2.8.1), either for all ℓ if Φ is positive definite, or for $\ell > L$ if Φ is conditionally positive definite with respect to Π_L.

3.1 Kernel-Based Quadrature

In practice the integrals involved in our Galerkin method will not be calculated directly. Instead, we will employ a *quadrature formula*, which is a weighted average of the values of the integrand at a denser set of centers.

Theorem 3.1.1. *Suppose Φ is conditionally positive definite with respect to Π_L, and that (2.8.1) holds for some $\tau \in \mathbb{N}$ with $\tau > d/2$, and all $\ell > L$. Suppose Ξ is a finite Π_L-unisolvent set of centers in $\mathbb{M}$. There is a unique quadrature formula*

$$\int_{\mathbb{M}} f d\mu \approx \int_{\mathbb{M}} I_{\Phi,\Xi,L} f d\mu = \sum_{\xi \in \Xi} w_\xi f(\xi) =: \mathcal{Q}^{\Xi}(f),$$

with weights $w = \{w_\xi\}_{\xi \in \Xi}$, that is exact on $V_{\Xi,L}$ and satisfies

$$\left| \int_{\mathbb{M}} f \; d\mu - \mathcal{Q}^{\Xi}(f) \right| \leq C h_{\Xi}^{\tau} \|f\|_{H^\tau}$$

for $f \in H^\tau$.

Proof. Let $\{p_1, \ldots, p_Q\}$ be a basis for Π_L. Let $\mathbf{1}$ be the $\#\Xi \times 1$ vector whose entries are all 1, and let

$$J_0 = \int_{\mathbb{M}} \Phi(\cdot, y) d\mu,$$

which is independent of which $y \in \mathbb{M}$ we choose. Also, let J be the $Q \times 1$ vector with entries

$$J_k = \int_{\mathbb{M}} p_k d\mu$$

for $k = 1, \ldots, Q$. We claim that if w and v are the (unique) $\#\Xi \times 1$ and $Q \times 1$ vectors that solve the system

$$K_{\Xi,L} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} J_0 \mathbf{1} \\ J \end{bmatrix},$$

then for all $s \in V_{\Xi,L}$,

$$\int_{\mathbb{M}} s \, d\mu = w^T \, s|_{\Xi} = \sum_{\xi \in \Xi} w_\xi s(\xi).$$

Indeed, if

$$s = \sum_{\xi \in \Xi} a_\xi \Phi(\cdot, \xi) + \sum_{k=1}^{Q} b_k p_k \in V_{\Xi,L},$$

then note that a and b satisfy the interpolation equation (2.4.2) with c replaced with $s|_{\Xi}$. Thus we have

$$\begin{aligned}
\int_{\mathbb{M}} s \, d\mu &= \sum_{\xi \in \Xi} a_\xi \int_{\mathbb{M}} \Phi(\cdot, \xi) d\mu + \sum_{k=1}^{Q} b_k \int_{\mathbb{M}} p_k d\mu \\
&= J_0 \mathbf{1}^T a + J^T b = \begin{bmatrix} J_0 \mathbf{1}^T & J^T \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \\
&= \begin{bmatrix} J_0 \mathbf{1}^T & J^T \end{bmatrix} K_{\Xi,L}^{-1} \begin{bmatrix} s|_{\Xi} \\ 0 \end{bmatrix} = \left(K_{\Xi,L}^{-1} \begin{bmatrix} J_0 \mathbf{1} \\ J \end{bmatrix} \right)^T \begin{bmatrix} s|_{\Xi} \\ 0 \end{bmatrix} \\
&= \begin{bmatrix} w \\ v \end{bmatrix}^T \begin{bmatrix} s|_{\Xi} \\ 0 \end{bmatrix} = w^T \, s|_{\Xi} \, .
\end{aligned}$$

This proves the exactness on $V_{\Xi,L}$. The uniqueness of the weights follows from the invertibility of $K_{\Xi,L}$. What remains to be shown is the error estimate. This follows directly from Hölder's inequality and the interpolation error estimate from Theorem 2.8.5 with $\sigma = 0$:

$$\begin{aligned}\left|\int_{\mathbb{M}} f\, d\mu - Q^{\Xi}(f)\right| &\le \int_{\mathbb{M}} |f - I_{\Phi,\Xi,L} f|\, d\mu \\ &\le (\mathrm{Vol}(\mathbb{M}))^{1/2} \|f - I_{\Phi,\Xi,L} f\|_{L_2} \le C h_{\Xi}^{\tau} \|f\|_{H^{\tau}}.\end{aligned}$$

□

There is a linear algebra technique, which can be found in [12, Section 2.2], that allows for numerically efficient computation of the weights w. They are given by $w = w_{\|} + w_{\perp}$, where $w_{\|} = P(P^T P)^{-1} J$ and $w_{\perp}$ satisfies the following system:

$$\begin{bmatrix} K_{\Xi} & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} w_{\perp} \\ v \end{bmatrix} = \begin{bmatrix} J_0 \mathbf{1} - A_{\Phi,\Xi} P(P^T P)^{-1} J \\ 0 \end{bmatrix}.$$

In fact, $w_{\perp}$ can be obtained without having to solve for v. See [12] for details.

3.2 The Lagrange Basis

There are obvious bases for our approximation spaces. If Φ is positive definite, then $\{\Phi(\cdot,\xi)\}_{\xi\in\Xi}$ is a basis for V_{Ξ}. If Φ is conditionally positive definite, then things are more complicated, because we can only allow certain linear combinations of these "translates", and we have to include a basis for Π_L. In both cases, however, there is an ideal basis to use: the Lagrange basis. For each $\xi \in \Xi$, we find the unique interpolant χ_ξ that interpolates $\delta_\xi = \{\delta_{\xi\eta}\}_{\eta\in\Xi}$, where $\delta_{\xi\eta} = 1$ if $\xi = \eta$ and $\delta_{\xi\eta} = 0$ otherwise. Thus, χ_ξ is the unique member of our approximation space for which $\chi_\xi(\eta) = \delta_{\xi\eta}$. The *Lagrange basis*, then, is $\{\chi_\xi\}_{\xi\in\Xi}$. Computing the coefficients $\alpha_\xi = \{\alpha_{\xi,\eta}\}_{\xi\in\Xi}$ and $\beta_\xi = \{\beta_{\xi,k}\}_{k\in\{1,\ldots,Q\}}$ of each Lagrange function

$$\chi_\xi = \sum_{\eta\in\Xi} \alpha_{\xi,\eta}\Phi(\cdot,\xi) + \sum_{k=1}^{Q} \beta_{\xi,k} p_k$$

can be computationally demanding; however, once computed interpolation becomes immediate:

$$I_{\Phi,\Xi,L} f = \sum_{\xi\in\Xi} f(\xi)\chi_\xi.$$

Of particular interest is the decay of the Lagrange functions. Bases that have rapid decay away from where they are centered are desirable because perturbations in data only affect interpolants locally. In the Euclidean setting, Madych and Nelson, in [21], considered the case where $\mathbb{M} = \mathbb{R}^d$ and $\Xi = \mathbb{Z}^d$. Their Lagrange functions, in turns out, are all translates of a single function L_k, which they referred to as "the fundamental function of interpolation for k-harmonic splines". By k-harmonic spline they mean, for $2k \geq d+1$, a function f for which $f \in C^{2k-d-1}$ and for which $\Delta^k f = 0$ on $\mathbb{R}^d \setminus \mathbb{Z}^d$. Here Δ is the usual Laplacian on $\mathbb{R}^d$. They showed that L_k decays exponentially in the distance from the origin. The kernel from Example 2.1.2 c is such a spline. More generally, for $m \in \mathbb{N}$ with $2m > d$, the *surface splines*

$$\Phi(x,y) = \begin{cases} \|x-y\|^{2m-d} & d \text{ odd}, \\ \|x-y\|^{2m-d}\log\|x-y\| & d \text{ even}, \end{cases}$$

generate Lagrange functions that decay exponentially.

Another author to investigate the cardinal setting is Buhmann, who showed in [4] that for a large number of families of radial basis functions, the *cardinal function* χ decays algebraically away from the origin. For example, if we use the kernel from Example 2.1.2(c), but on $\mathbb{R}^3$ instead of on $\mathbb{R}^2$, we obtain $|\chi_\xi(x)| \leq C\|x\|_2^{-4}$, but no faster. Another way of putting it is that $\lim_{\|x\|_2\to\infty} |\chi_\xi(x)|\|x\|_2^4$ is finite, but nonzero. Yet another way is that $\chi_\xi(x) = O\left(\|x\|_2^{-4}\right)$ but $\chi_\xi(x) \neq o\left(\|x\|_2^{-4}\right)$ as $\|x\|_2 \to \infty$. Again, the Lagrange functions in this setting are all translates of a single function.

If Φ is positive definite and satisfies (2.8.1) with $\tau = m \in \mathbb{N}$ with $\tau > d/2$, Proposition 2.4.4 gives that χ_ξ is the unique $\mathcal{N}_\Phi$ norm-minimizing element u of $\mathcal{N}_\Phi$ such that $u|_\Xi = \delta_\xi$. We can therefore compare χ_ξ to a bump function Ψ. We do this first on Euclidean space, and then we will use the metric equivalence from Lemma 2.6.1 to transfer the result to $\mathbb{M}$. Let $\psi : [0,\infty) \to [0,1]$ satisfy $\psi \in C^\infty([0,\infty))$, $\text{supp}\psi = [0,1)$, and $\psi(0) = 1$, and define $\Psi : \mathbb{R}^d \to [0,1]$ by setting $\Psi(x) = \psi(\|x\|_2/q_\Xi)$. Clearly $\text{supp}\Psi = B(0, q_\Xi)$, $\Psi \in W_2^m(\mathbb{R}^d)$, and $\Psi|_\Xi = \delta_\xi$, and it is routine to show that $\|\Psi\|_{W_2^m(\mathbb{R}^d)} \leq C q_\xi^{\frac{d}{2}-m}$. Hence,

$$\|\chi_\xi\|_{W_2^m} \leq C q_\Xi^{\frac{d}{2}-m}, \tag{3.2.1}$$

which we refer to from here on as our *bump estimate*.

We consider two categories of kernels: those whose Lagrange functions decay algebraically away from their center, and those whose Lagrange functions decay exponentially away from their center. These two types of decay can both be encapsulated in what are called *energy estimates*. In the sequel, we will denote Euclidean balls of radius r centered at $x \in \mathbb{R}^d$ by $B(x,r)$, and balls on the manifold of radius r centered at $p \in \mathbb{M}$ by $b(p,r)$.

Definition 3.2.2. Suppose Φ is a kernel which satisfies (2.8.1) for $\tau = m \in \mathbb{N}$ with $m > d/2$, for all $\ell \in \mathbb{N}$ if Φ is positive definite or for all $\ell > L \in \mathbb{N}$ if Φ is conditionally positive definite with respect to Π_L.

We say that Φ provides an *algebraic energy estimate* if there are positive constants C such that for a set of centers Ξ, the Lagrange basis $\{\chi_\xi\}_{\xi\in\Xi}$

satisfies, for h_Ξ sufficiently small and $0 \le r \le \mathrm{diam}(\mathbb{M})$,

$$\|\chi_\xi\|_{W_2^m(b(\xi,r)^c)} \le C q_\Xi^{\frac{d}{2}-m} \left(1 + \frac{r}{h_\Xi}\right)^{-2m}. \tag{3.2.3}$$

We say Φ provides an *exponential energy estimate* if there are positive constants C and ν such that for a set of centers Ξ, the Lagrange basis $\{\chi_\xi\}_{\xi\in\Xi}$ satisfies, for h_Ξ sufficiently small and $0 \le r \le \mathrm{diam}(\mathbb{M})$,

$$\|\chi_\xi\|_{W_2^m(b(\xi,r)^c)} \le C q_\Xi^{\frac{d}{2}-m} \exp\left(-\nu \frac{r}{h_\Xi}\right). \tag{3.2.4}$$

We note that both energy estimates with $r = 0$ agree with our bump estimate. Energy estimates will allow us to obtain pointwise bounds on $\chi_\xi(x)$ and $\nabla\chi_\xi(x)$ that depend only on the distance between x and ξ, but first, let us elaborate on the situation described in section 2.3, where we considered kernels whose coefficients were reciprocals of polynomials in λ_ℓ. These types of kernels, it turns out, provide energy estimates.

Definition 3.2.5. Let $m \in \mathbb{N}$ with $m > d/2$. We call $\Phi_m : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ *polyharmonic* if it is of the form (2.2.1) with coefficients $\widehat{\phi}_m(\ell) = Q(\lambda_\ell)^{-1}$, where $Q(x) = \sum_{k=0}^m c_k x^k$ is a polynomial of degree m with $c_m > 0$. Corresponding to a polyharmonic kernel Φ_m is a differential operator $\mathcal{L}_m = Q(\Delta)$ of order $2m$ for which Φ_m is the fundamental solution. When we write that Φ_m is polyharmonic, it is understood that the degree of the polynomial Q is captured in the index m.

Note that the condition $c_m > 0$ means that $\widehat{\phi}(\ell) = Q(\lambda_\ell)^{-1} > 0$ for all ℓ larger than some $L \in \mathbb{N}$. Hence, Φ_m is conditionally positive definite with respect to Π_L. The results in [16, Section 5] say that if, in addition, $\mathcal{L}_m$ annihilates the auxiliary space Π_L, then Φ_m provides an exponential energy estimate. Obviously, then, if $\widehat{\phi}_m(\ell) > 0$ for all $\ell \in \mathbb{N}$, then Φ_m is positive definite and thus also provides an exponential energy estimate, since in this case the auxiliary space is trivial. If Φ_m is conditionally positive definite with respect to Π_L and $\mathcal{L}_m$ does *not* annihilate the auxiliary space Π_L, then it is only known that Φ_m provides an algebraic energy estimate. As the authors in [16] mention, while only an algebraic energy estimate has been proven in this case, there is sufficient cause to believe that this is just an artifact of the proof. In any case, we will examine both possibilities, algebraic and exponential.

Note also that if Φ_m is polyharmonic, then the coefficients $\widehat{\phi}_m(\ell) = Q(\lambda_\ell)^{-1}$ satisfy (2.8.1) with $\tau = m$ and $\ell > L$, since both $Q(\lambda_\ell)^{-1}$ and $(1 + \lambda_\ell)^{-m}$ are reciprocals of polynomials in λ_ℓ of degree m with positive leading coefficients.

To recapitulate, if $\mathbb{M}$ is two-point homogeneous and Φ_m is polyharmonic, then Φ_m provides an energy estimate and $\mathcal{N}_{\Phi_m} = H^m$. In general, this is not the case. There are also kernels Φ_m that provide energy estimates with parameter m that are not polyharmonic. This thesis treats those as well.

We note that our definition of a kernel providing an energy estimate includes the assumption that (2.8.1) holds for $t = m \in \mathbb{N}$, at least for ℓ large enough, and that $m > d/2$. As such, $\mathcal{N}_{\Phi_m} = H^m$, and we will not repeat the assumption $m > d/2$.

Proposition 3.2.6. *Suppose Φ_m provides an energy estimate, Ξ is a finite subset of $\mathbb{M}$ (Π_L-unisolvent if Φ_m is conditionally positive definite with respect to Π_L), and $\{\chi_\xi\}_{\xi \in \Xi}$ is the corresponding Lagrange basis.*

If Φ_m provides an algebraic energy estimate as in (3.2.3), then for h_Ξ sufficiently small,

$$|\chi_\xi(p)| \leq C \left(1 + \frac{\operatorname{dist}(p, \xi)}{h_\Xi}\right)^{-2m}.$$

If Φ provides an exponential energy estimate as in (3.2.4), then for h_Ξ sufficiently small,

$$|\chi_\xi(p)| \le C\exp\left(-\nu\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right).$$

Proof. Let $r = \operatorname{dist}(p,\xi)$. First, we have $|\chi_\xi(p)| \le \|\chi_\xi\|_{L_\infty(b(\xi,r)^c)}$. We now use a Zeros Lemma, but not the one from Lemma 2.7.1, which only applies to the whole space. We use instead the Zeros Lemma from [16, Theorem A.11], which applies to complements of balls. Thus,

$$|\chi_\xi(p)| \le Ch_\Xi^{m-\frac{d}{2}} \|\chi_\xi\|_{W_2^m(b(\xi,r)^c)},$$

and the results now follow from the energy estimates. □

A similar proof, but using [16, Corollary A.15] instead of [16, Theorem A.11], gives the following.

Proposition 3.2.7. *Suppose Φ_m provides an energy estimate, Ξ is a finite subset of $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and $\{\chi_\xi\}_{\xi\in\Xi}$ is the corresponding Lagrange basis. If Φ provides an algebraic energy estimate as in (3.2.3), then for h_Ξ sufficiently small,*

$$|\nabla\chi_\xi(p)| \le Ch_\Xi^{-1}\left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m}.$$

If Φ provides an exponential energy estimate as in (3.2.4), then

$$|\nabla\chi_\xi(p)| \le Ch_\Xi^{-1}\exp\left(-\nu\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right).$$

The decay of the Lagrange functions has been used to prove the following stability bound, which we refer to as *the L_2-stability of the Lagrange basis.* It says that the ℓ_2 norm of a vector can be controlled by the L_2 norm of the corresponding linear combination of Lagrange functions, and vice versa.

Theorem 3.2.8. ([16, Theorem 5.7]) *Suppose Φ_m provides an energy estimate, Ξ is a finite subset of $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and $\{\chi_\xi\}_{\xi\in\Xi}$ is the corresponding Lagrange basis. There exist constants C_1 and C_2 such that, for h_Ξ sufficiently small,*

$$C_1 h_\Xi^{d/p} \|\beta\|_{\ell_2(\Xi)} \leq \left\| \sum_{\xi\in\Xi} \beta_\xi \chi_\xi \right\|_{L_2(\mathbb{M})} \leq C_2 h_\Xi^{d/p} \|\beta\|_{\ell_2(\Xi)}$$

holds for all $\beta = \{\beta_\xi\}_{\xi\in\Xi} \in \ell_2(\Xi)$.

We end this section by noting a connection between the Lagrange functions and our quadrature rule; namely, that the weights are the integrals of the Lagrange functions. This is because χ_ξ is in the approximation space, and hence the quadrature rule is exact:

$$\int_{\mathbb{M}} \chi_\xi d\mu = w^T \left.\chi_\xi\right|_\Xi = w^T \delta_\xi = w_\xi.$$

While this is a useful theoretical fact, and one that we will use later when showing the off-diagonal decay of our quadratized stiffness matrix, it is important to stress that this has *nothing* to do with how the weights are actually computed. They are computed using the linear algebra technique discussed at the end of section 3.1

3.3 Kernel-Based Galerkin Methods

We will be approximating weak solutions to a differential equation of the form $\mathcal{L}u = f$, with $\mathcal{L}$ given in *divergence form* by

$$\mathcal{L}u = -\nabla^* a\left(\nabla u, \cdot\right) + bu. \tag{3.3.1}$$

Let us first be more precise about what this means. We start with a symmetric rank $(1,1)$ tensor a. This means that for each $p \in \mathbb{M}$,

$$a(p) : T_p^*\mathbb{M} \times T_p\mathbb{M} \longrightarrow \mathbb{R}.$$

If $f \in C^\infty(\mathbb{M})$, and $p \in \mathbb{M}$, then $\nabla f(p)$ is a rank 1 covariant tensor; i.e. $\nabla f(p) : T_p\mathbb{M} \to \mathbb{R}$, $v \mapsto \nabla_v f(p)$, which is a directional derivative. If we put this as the first argument in $a(p)$ and leave the second open, we get another rank 1 covariant tensor,

$$a(p)\left(\nabla f(p), \cdot\right) : T_p\mathbb{M} \longrightarrow \mathbb{R}.$$

Being a rank 1 covariant tensor, we can apply ∇^* to it to obtain a scalar, $\nabla^* a(\nabla f, \cdot)$. This is what is meant in the principle part of (3.3.1), the coordinate-free expression of our differential equation.

In coordinates $(\mathfrak{U}, \phi)$, where $\phi(\mathfrak{U}) = U \subset \mathbb{R}^d$ and $\phi(p) = (x^1, \ldots, x^d) \in U$, let $\mathbf{e}_1, \ldots, \mathbf{e}_d$ be a basis for $T_p\mathbb{M}$ and $\mathbf{e}^1, \ldots, \mathbf{e}^d$ the dual basis for $T_p^*\mathbb{M}$. Then $\nabla f(p) = \sum_j \frac{\partial f}{\partial x^j}\mathbf{e}^j$ and $a(p) = \sum_{j,k} a_j^k \mathbf{e}_k \otimes \mathbf{e}^j$, and so

$$a(p)\left(\nabla f(p), \cdot\right) = \sum_{j,k} a_j^k \mathbf{e}_k \otimes \mathbf{e}^j \left(\sum_\ell \frac{\partial f}{\partial x^\ell}\mathbf{e}^\ell, \cdot \right) = \sum_{j,k} a_j^k \frac{\partial f}{\partial x^k}\mathbf{e}^j.$$

By the definition of ∇^*, for any covariant rank 1 tensor $\mathbf{T}$,

$$\int_\mathbb{M} \left\langle \nabla f, \mathbf{T} \right\rangle_p d\mu(p) = \int_\mathbb{M} f(p)\nabla^*\mathbf{T}(p) d\mu(p).$$

Recall that if $\mathbf{S} = \sum_j S_j \mathbf{e}^j$ and $\mathbf{T} = \sum_k T_k \mathbf{e}^k$, then

$$\left\langle \mathbf{S}, \mathbf{T} \right\rangle_p = \sum_{j,k} S_j T_k \left\langle \mathbf{e}^j, \mathbf{e}^k \right\rangle_p = \sum_{j,k} S_j T_k g^{jk}.$$

So, in a coordinate neighborhood $\mathfrak{U}$,

$$\int_{\mathfrak{U}} \langle \mathbf{S}, \mathbf{T} \rangle_p \, d\mu(p) = \int_U \sum_{j,k} S_j T_k g^{jk}(x) \sqrt{\det(g(x))} dx.$$

Replacing S with ∇f, and assuming that the closure of the support of f is contained in $\mathfrak{U}$, we have

$$\begin{aligned}\int_{\mathfrak{U}} \langle \nabla f, \mathbf{T} \rangle_p \, d\mu(p) &= \int_U \sum_{j,k} \frac{\partial f}{\partial x^j} T_k g^{jk}(x) \sqrt{\det(g(x))} dx \\ &= -\int_U f(\phi(x)) \sum_{j,k} \frac{\partial}{\partial x^j} \left[T_k g^{jk}(x) \sqrt{\det(g(x))} \right] dx.\end{aligned}$$

Multiplying and dividing by $\sqrt{\det(g(x))}$ gives

$$\begin{aligned}&\int_{\mathfrak{U}} f(p) \nabla^* \mathbf{T}(p) d\mu(p) \\ &= -\int_U f(\phi(x)) \frac{1}{\sqrt{\det(g(x))}} \sum_{j,k} \frac{\partial}{\partial x^j} \left[T_k g^{jk}(x) \sqrt{\det(g(x))} \right] \sqrt{\det(g(x))} dx.\end{aligned}$$

This shows that, in coordinates,

$$\nabla^* \mathbf{T} = -\frac{1}{\sqrt{\det(g(x))}} \sum_{j,k} \frac{\partial}{\partial x^j} \left[T_k g^{jk} \sqrt{\det(g(x))} \right].$$

If $\mathbf{T} = a(p) (\nabla f(p), \cdot)$, then $T_k = \sum_\ell a_k^l \frac{\partial f}{\partial x^\ell}$, and so, in coordinates, our differential equation is

$$\mathcal{L}u = \frac{1}{\sqrt{\det(g)}} \sum_{j,k,\ell=1}^{d} \frac{\partial}{\partial x^j} \left[g^{jk} \sqrt{\det(g)} a_k^\ell \frac{\partial u}{\partial x^\ell} \right] + bu = f. \qquad (3.3.2)$$

If we multiply the principle part of (3.3.1) by $v \in H^1$ and integrate, we get

$$\int_{\mathbb{M}} v(p)\left(\nabla^* a(p)\left(\nabla u(p), \cdot\right)\right) d\mu(p) = \int_{\mathbb{M}} \left\langle a\left(\nabla u, \cdot\right), \nabla v\right\rangle_p d\mu(p)$$
$$= \int_{\mathbb{M}} a^\sharp(p)\left(\nabla u(p), \nabla v(p)\right) d\mu(p).$$

Thus, the weak form of (3.3.1) is

$$\langle u, v\rangle_{a,b} := \int_{\mathbb{M}} \left(a^\sharp\left(\nabla u, \nabla v\right) + buv\right) d\mu = \int_{\mathbb{M}} fv \; d\mu =: \lambda_f(v). \qquad (3.3.3)$$

We note three things. First, that the components of $a^\sharp$ are $a^{ij} = \sum_k g^{jk} a^i_k$. Next, in the case $a^\flat = g$ and $b = 0$ the equation reduces to $-\Delta u = f$. Finally, $\langle u, v\rangle_{a,b} = \langle \mathcal{L}u, v\rangle_{L_2}$.

We assume there are constants $0 < b_1 \leq b_2$ such that $b_1 \leq b(p) \leq b_2$ for all $p \in \mathbb{M}$. We also assume that $a^\sharp$ is symmetric and positive definite in the sense that there are constants $0 < a_1 \leq a_2$ such that

$$a_1\langle v, v\rangle_p \leq a^\sharp(p)(v, v) \leq a_2\langle v, v\rangle_p$$

for each $p \in \mathbb{M}$ and $v \in T^*_p\mathbb{M}$. Finally, we assume that $a^\sharp$ is bounded in the sense that

$$\left|a^\sharp\left(v, w\right)\right| \leq C\left|v\right|\left|w\right|$$

for all $v, w \in T^*_p\mathbb{M}$.

The conditions on a and b ensure that $\langle\cdot,\cdot\rangle_{a,b}$ is bounded and coercive, and λ_f is bounded. Hence, the Lax-Milgrim Theorem applies, and we have the following.

Proposition 3.3.4. *The bilinear form $\langle\cdot,\cdot\rangle_{a,b}$ is coercive and bounded on H^1 and defines an inner product on H^1, and the norms $\|\cdot\|_{a,b}$ and $\|\cdot\|_{H^1}$*

are equivalent. In addition, if $f \in L_2$ *there is a unique* $u \in H^1$ *such that* $\langle u, v\rangle_{a,b} = \lambda_f(v)$ *for all* $v \in H^1$.

The next proposition holds on coordinate patches by [9, pp. 261-269]. Since $\mathbb{M}$ is compact, it holds globally.

Proposition 3.3.5. *If* u *is the weak solution to* $\mathcal{L}u = f$*, with* $f \in H^s$*,* $s \geq 0$*, then there is a constant* C *that is independent of* u *and* f *such that* $\|u\|_{H^{s+2}} \leq C\|f\|_{H^s}$.

The *weak solution* u to $\mathcal{L}u = f$ is the solution to the following problem: given $f \in L_2$, find $u \in H^1$ such that $\langle u, v\rangle_{a,b} = \lambda_f(v)$ for all $v \in H^1$. The Galerkin approximation scheme modifies this problem to read: given $f \in L_2$ find $u_\Xi \in V_{\Xi,L}$ such that $\langle u_\Xi, v\rangle_{a,b} = \lambda_f(v)$ for all $v \in V_{\Xi,L}$. The *Galerkin approximation* u_Ξ to the solution u of $\mathcal{L}u = f$ is obtained as follows. The *stiffness matrix* in the Lagrange basis, $B_\Xi = \{B_{\xi\eta}\}_{\xi,\eta\in\Xi}$, has entries

$$B_{\xi\eta} = \langle \chi_\xi, \chi_\eta\rangle_{a,b}$$

The Galerkin approximation, then, is

$$u_\Xi = \sum_{\xi\in\Xi} \gamma_\xi \chi_\xi,$$

where the coefficients $\gamma = \{\gamma_\xi\}_{\xi\in\Xi}$ are obtained by solving the linear system $B_\Xi\gamma = \omega$, with $\omega = \{\omega_\xi\}_{\xi\in\Xi}$ the vector with entries $\omega_\xi = \lambda_f(\chi_\xi)$.

By construction, $\langle u_\Xi, \chi_\xi\rangle_{a,b} = \lambda_f(\chi_\xi)$ for each $\xi \in \Xi$; thus, $\langle u_\Xi, v\rangle_{a,b} = \lambda_f(v)$ for each $v \in V_{\Xi,L}$. Since u is the weak solution to $\mathcal{L}u = f$, surely $\langle u, v\rangle_{a,b} = \lambda_f(v)$. Hence $\langle u - u_\Xi, v\rangle_{a,b} = 0$ for all $v \in V_{\Xi,L}$, and so we see that u_Ξ is the orthogonal projection of u onto $V_{\Xi,L}$ in the inner product $\langle\cdot,\cdot\rangle_{a,b}$.

We now give an error estimate for Galerkin approximation. We state it in the case of a conditionally positive definite kernel. Since a positive definite kernel is conditionally positive definite with respect to any auxiliary space, it holds for positive definite kernels as well.

Theorem 3.3.6. *Suppose Φ_m provides an energy estimate and Ξ is a finite subset of $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive with respect to Π). Let u_Ξ be the corresponding Galerkin approximation to the solution of $\mathcal{L}u = f$, where $f \in H^{m-2}$. There exists a positive constant C that is independent of f and Ξ such that, for h_Ξ sufficiently small,*

$$\|u - u_\Xi\|_{H^1} \le C h_\Xi^{m-1} \|f\|_{H^{m-2}} .$$

Proof. By the equivalence of the norms $\|\cdot\|_{H^1}$ and $\|\cdot\|_{a,b}$, and because u_Ξ is the orthogonal projection of u onto $V_{\Xi,L}$ with respect to the inner product $\langle\cdot,\cdot\rangle_{a,b}$,

$$\|u - u_\Xi\|_{H^1} \le C \|u - u_\Xi\|_{a,b} = C \min_{v \in V_{\Xi,L}} \|u - v\|_{a,b} \le C \|u - I_{\Xi,L} u\|_{a,b} .$$

Using that norm equivalence again and Theorem 2.8.5 with $\tau = m$ and $\sigma = 1$ gives

$$\|u - u_\Xi\|_{H^1} \le C \|u - I_{\Xi,L} u\|_{H^1} \le C h_\Xi^{m-1} \|u\|_{H^m} .$$

Finally, by Proposition 3.3.5,

$$\|u - u_\Xi\|_{H^1} \le C h_\Xi^{m-1} \|f\|_{H^{m-2}} .$$

□

3.4 The Stiffness Matrix in the Lagrange Basis

It is not necessary to use the Lagrange basis to compute the Galerkin approximation; however, one of the advantages is that the entries of the stiffness matrix decay away from the main diagonal. This makes computing the Galerkin approximation numerically stable. We illustrate this decay now.

Lemma 3.4.1. *Suppose Φ_m provides an energy estimate, Ξ is a finite set of centers in $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and B_Ξ is the stiffness matrix in the Lagrange basis for Galerkin approximation using the kernel Φ_m and centers Ξ.*

If Φ_m provides an algebraic energy estimate as in (3.2.3), then there is a positive constant C depending only on Φ_m and ρ such that, for h_Ξ sufficiently small,

$$|B_{\xi\eta}| \leq C h_\Xi^{d-2} \left(1 + \frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right)^{-2m}.$$

If Φ_m provides an exponential energy estimate as in (3.2.4), then there is a positive constant C depending only on Φ_m and ρ such that, for h_Ξ sufficiently small,

$$|B_{\xi\eta}| \leq C h_\Xi^{d-2} \exp\left(-\nu \frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right).$$

Proof. First suppose Φ_m provides an algebraic energy estimate. By Proposition 3.2.6,

$$\left|\int_{\mathbb{M}} b\chi_\xi\chi_\eta d\mu\right| \le Cb_2\int_{\mathbb{M}}\left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m}\left(1+\frac{\operatorname{dist}(p,\eta)}{h_\Xi}\right)^{-2m} d\mu(p). \tag{3.4.2}$$

We show a similar bound holds for the principle part. Since $\mathbb{M}$ is compact, the open cover $\{b(q, r_{\mathbb{M}}/3)\}_{q\in\mathbb{M}}$ has a finite subcover $\left\{b(q_i, r_{\mathbb{M}/3})\right\}_{i=1}^n$. Let $\Omega_1 = b(q_1, r_{\mathbb{M}}/3)$ and for $i = 2, \ldots, n$ set

$$\Omega_i = b(q_i, r_{\mathbb{M}}/3) \setminus \bigcup_{j=1}^{i-1} \Omega_j.$$

Then $\mathbb{M} = \bigcup_{i=1}^n \Omega_i$ is a disjoint union. Note that each $\operatorname{Exp}_i = \operatorname{Exp}_{q_i}$ is a bijective isometry from Ω_i to $U_i = \operatorname{Exp}_i^{-1}(\Omega_i)$, and so

$$\begin{aligned}\left|\int_{\mathbb{M}} a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right) d\mu\right| &= \left|\sum_{j=1}^n \int_{\Omega_i} a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right) d\mu\right| \\ &= \left|\sum_{i=1}^n \int_{U_i} \sum_{j,k} a^{jk}(x)\frac{\partial\chi_\xi}{\partial x^j}(x)\frac{\partial\chi_\eta}{\partial x^k}(x)\ dx\right| \\ &\le C\sum_{i,j,k}\int_{U_i}\left|\frac{\partial\chi_\xi}{\partial x^j}(x)\right|\left|\frac{\partial\chi_\eta}{\partial x^k}(x)\right| dx,\end{aligned}$$

where we have abused notation slightly by writing, for example, $\frac{\partial\chi_\xi}{\partial x^j}(x)$ instead of $\frac{\partial\chi_\xi\circ\operatorname{Exp}_i}{\partial x^j}(x)$. Surely $\left|\frac{\partial\chi_\xi}{\partial x^j}(x)\right| \le |\nabla\chi_\xi(x)|$, and so by Theorem 3.2.7,

$$\begin{aligned}&\left|\int_{\mathbb{M}} a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right) d\mu\right| \\ &\le Ch_\Xi^{-2}\int_{\mathbb{M}}\left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m}\left(1+\frac{\operatorname{dist}(p,\eta)}{h_\Xi}\right)^{-2m} d\mu(p).\end{aligned} \tag{3.4.3}$$

We now split up the integrals in (3.4.2) and (3.4.3) onto two half spaces.

Let

$$\Omega_\xi = \{x \in \mathbb{M} : \operatorname{dist}(x,\xi) < \operatorname{dist}(x,\eta)\}$$

and

$$\Omega_\eta = \{x \in \mathbb{M} : \operatorname{dist}(x,\eta) < \operatorname{dist}(x,\xi)\}.$$

Then $\Omega_\xi \cup \Omega_\eta$ is, minus a set of measure zero, $\mathbb{M}$, and $\Omega_\xi \cap \Omega_\eta = \emptyset$. Hence, for an integrable function f,

$$\int_{\mathbb{M}} f \, d\mu = \int_{\Omega_\xi} f \, d\mu + \int_{\Omega_\eta} f \, d\mu.$$

Note that for $x \in \Omega_\xi$, $\operatorname{dist}(x,\eta) \geq \frac{1}{2}\operatorname{dist}(\xi,\eta)$, and so

$$\begin{aligned}
&\int_{\Omega_\xi} \left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} \left(1 + \frac{\operatorname{dist}(p,\eta)}{h_\Xi}\right)^{-2m} d\mu(p) \\
&\leq C \left(1 + \frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right)^{-2m} \int_{\mathbb{M}} \left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} d\mu(p).
\end{aligned}$$

Now,

$$\begin{aligned}
&\int_{\mathbb{M}} \left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} d\mu(p) \\
&= \int_{b(\xi,r_{\mathbb{M}})} \left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\xi}\right)^{-2m} d\mu(p) \\
&\qquad\qquad + \int_{b(\xi,r_{\mathbb{M}})^c} \left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} d\mu(p).
\end{aligned}$$

For $p \in \mathbb{M}$ with $\operatorname{dist}(p,\xi) \geq r_{\mathbb{M}}$,

$$\left(1 + \frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} \leq r_{\mathbb{M}}^{-2m} h_\Xi^{2m},$$

and so

$$\int_{b(\xi,r_{\mathbb{M}})^c} \left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} d\mu(p) \le \operatorname{Vol}(\mathbb{M}) r_{\mathbb{M}}^{-2m} h_\Xi^{2m} \le C h_\Xi^d.$$

Lemma 2.6.1 applied to $g(p) = \left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-m}$ with $j = 0$ and $\Omega = \operatorname{Exp}_\xi^{-1}(b(\xi, r_{\mathbb{M}}))$ gives

$$\begin{aligned}
&\int_{b(\xi,r_{\mathbb{M}})} \left(1+\frac{\operatorname{dist}(p,\xi)}{h_\Xi}\right)^{-2m} d\mu(p) = \|g\|^2_{L_2(b(\xi,r_{\mathbb{M}}))} \\
&\le C \left\| g \circ \operatorname{Exp} \right\|^2_{L_2(\operatorname{Exp}_\xi^{-1}(b(\xi,r_{\mathbb{M}})))} \\
&\le C \int_{\mathbb{R}^d} \left(1+\frac{\|x\|_2}{h_\Xi}\right)^{-2m} dx = C' h_\Xi^d,
\end{aligned}$$

where we have used the integrability of $(1+\|x\|_2)^{-2m}$. The same arguments apply to the integral over Ω_η, and therefore

$$|B_{\xi\eta}| \le C' h_\Xi^{d-2} \left(1+\frac{\operatorname{dist}(\xi,\eta)}{h_\xi}\right)^{-2m}.$$

The result in the case that Φ_m provides an exponential energy estimate is proved in an almost identical way. □

Lemma 3.4.4. *Suppose Φ_m provides an energy estimate, Ξ is a finite set of centers in $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and $\{\chi_\xi\}_{\xi\in\Xi}$ is the corresponding Lagrange basis. Let B_Ξ be the stiffness matrix in the Lagrange basis. There is a positive constant C depending only on Φ_m and ρ such that, for h_Ξ sufficiently small,*

$$\left\| B_\Xi^{-1} \right\|_2 \le C h_\Xi^{-d}.$$

Proof. It suffices to show that $\lambda_{\min}(B_\Xi) \geq C q_\Xi^d$, for then $\left\| B_\Xi^{-1} \right\|_2 \leq C q_\Xi^{-d} = C\rho^d h_\Xi^{-d}$. For that it suffices to show that $v^T B_\Xi v \geq C q_\Xi^d \|v\|_{\ell_2(\Xi)}^2$ for all vectors $v = \{v_\xi\}_{\xi\in\Xi}$. Given such a v, let $u = \sum_{\xi\in\Xi} v_\xi \chi_\xi$, so that $\nabla u = \sum_{\xi\in\Xi} v_\xi \nabla\chi_\xi$. We have

$$
\begin{aligned}
v^T B_\Xi v &= \sum_{\xi\in\Xi}\sum_{\eta\in\Xi} v_\xi v_\eta \int_{\mathbb{M}} \left(a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right) + b\chi_\xi\chi_\eta\right) d\mu \\
&= \int_{\mathbb{M}} \left(a^\sharp \left(\sum_{\xi\in\Xi} v_\xi \nabla\chi_\xi, \sum_{\eta\in\Xi} v_\eta \nabla\chi_\eta \right) + b \left(\sum_{\xi\in\Xi} v_\xi\chi_\xi \right) \left(\sum_{\eta\in\Xi} v_\eta \chi_\eta \right) \right) d\mu \\
&= \int_{\mathbb{M}} \left(a^\sharp\left(\nabla u, \nabla u\right) + bu^2\right) d\mu.
\end{aligned}
$$

The metric tensor a being positive definite ensures $a^\sharp(\nabla u, \nabla u) > 0$. Hence we can use the L_2 stability of the Lagrange basis to obtain

$$
v^T B_\Xi v \geq \int_{\mathbb{M}} bu^2 d\mu \geq b_1 \|u\|_{L_2}^2 = b_1 \left\| \sum_{\xi\in\Xi} v_\xi \chi_\xi \right\|_{L_2}^2 \geq C_1^2 b_1 q_\Xi^d \|v\|_{\ell_2(\Xi)}^2 .
$$

$\square$

Chapter 4

Quadratized Kernel-Based Galerkin Approximation

In this chapter we explore what happens when we replace every integral in our Galerkin approximation scheme with the corresponding quadrature approximation. In Section 4.1 we introduce the quadratized stiffness matrix and establish bounds on the entries and norms of the quadratized stiffness matrix and the difference between the stiffness matrix and the quadratized stiffness matrix. In Section 4.2 we prove an error estimate between the Galerkin approximation and the quadratized Galerkin approximation in Theorem 4.2.1, which immediately translates into an error estimate between the weak solution and the quadratized Galerkin approximation in Corollary 4.2.7. We end this chapter by detailing algorithms for computing the quadratized stiffness matrix.

4.1 The Quadratized Stiffness Matrix

The *quadratized stiffness matrix* is the matrix obtained from the stiffness matrix by replacing each entry, which is an integral, by a quadrature estimate of that integral. The kernel used for quadrature is the same, but the centers used for quadrature are denser than the centers used for the Galerkin approximation.

Suppose Φ_m and Ξ are the kernel and centers, respectively, used for Galerkin approximation. Let Λ be centers used for quadrature. The quadratized stiffness matrix, B_{Ξ}^{Λ}, has entries $B_{\xi\eta}^{\Lambda} = Q^{\Lambda}\left(a^{\sharp}\left(\nabla\chi_{\xi}, \nabla\chi_{\eta}\right) + b\chi_{\xi}\chi_{\eta}\right)$. We make further assumptions on m, Ξ, and Λ, which are listed in Lemma 4.1.1.

The main objective of this section, Lemmas 4.1.16 and 4.1.21, is to bound the 2-norm of the difference between the stiffness matrix and the quadratized stiffness matrix. To that end we will need the following propositions. Lemma 4.1.1 gives a bound on the entries of the quadratized stiffness matrix as a function of their distance from the diagonal. Theorem 4.1.5 gives a bound on the Sobolev norms of the principle part and the non-principle part of the weak form of the equation $\mathcal{L}u = f$ when the inputs are Lagrange functions. Both are used after quadrature is applied when quadratizing the stiffness matrix. The theorem is modeled after [22, Theorem 4.4], though the differences in the proof are significant. Lemma 4.1.14 gives a bound on the entries of the difference between the stiffness matrix and the quadratized stiffness matrix. Lemmas 4.1.16 and 4.1.21 bound the 2-norm of that difference. We then have Lemma 4.1.25, which controls the 2-norm of the inverse of the quadratized stiffness matrix. It includes a threshold upper bound for h_{Λ} in terms of h_{Ξ} in order to work. As will be discussed, there is an alternative way to achieve the result in

Lemma 4.1.25 if we know that the quadrature weights are positive. This is done in Lemma 4.2.12.

In what follows whenever we say "there is a positive constant C" it is understood that C depends on Φ_m, ρ, and the parameters a and b from our differential equation (3.3.1). If other dependencies are specified, those are in addition to those just mentioned. We also note that the condition $m > d/2$ is included in Definition 3.2.2, and as such will not be repeated.

Lemma 4.1.1. *Suppose Φ_m provides an energy estimate, Ξ is a finite set of centers in $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and $\{\chi_\xi\}_{\xi\in\Xi}$ is the corresponding Lagrange basis. Let Λ be another set of centers in $\mathbb{M}$. Assume $m \geq 2$, $\#\Xi < \#\Lambda$, and $h_\Lambda < h_\Xi$. Let B_Ξ^Λ be the quadratized stiffness matrix.*

If Φ_m provides an algebraic energy estimate as in (3.2.3), then there is a positive constant C such that, for h_Ξ sufficiently small,

$$\left|B_{\xi\eta}^\Lambda\right| \leq C h_\Xi^{d-2} \left(1 + \frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right)^{-2m}.$$

If Φ_m provides an exponential energy estimate as in (3.2.4), then there is a constant C such that, for h_Ξ sufficiently small,

$$\left|B_{\xi\eta}^\Lambda\right| \leq C h_\Xi^{d-2} \exp\left(-\nu \frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right).$$

Proof. First, we have by Theorems 3.2.6 and 3.2.7 that

$$\begin{aligned}\left|B_{\xi\eta}^{\Lambda}\right| &= \left|\sum_{\lambda\in\Lambda}\left(a^{\sharp}(\lambda)\left(\nabla\chi_{\xi},\nabla\chi_{\eta}\right)+b(\lambda)\chi_{\xi}(\lambda)\chi_{\eta}(\lambda)\right)w_{\lambda}\right| \\ &\le C\sum_{\lambda\in\Lambda}\left(\left|\nabla\chi_{\xi}(\lambda)\right|\left|\nabla\chi_{\eta}(\lambda)\right|+\left|\chi_{\xi}(\lambda)\right|\left|\chi_{\eta}(\lambda)\right|\right)\left|w_{\lambda}\right| \\ &\le Ch_{\Xi}^{-2}\sum_{\lambda\in\Lambda}\left(1+\frac{\operatorname{dist}(\lambda,\xi)}{h_{\Xi}}\right)^{-2m}\left(1+\frac{\operatorname{dist}(\lambda,\eta)}{h_{\Xi}}\right)^{-2m}\left|w_{\lambda}\right|.\end{aligned}$$

Let $\{\widetilde{\chi}_{\lambda}\}_{\lambda\in\Lambda}$ be the Lagrange basis for Φ_m and Λ. By the L_1 stability of this basis,

$$|w_{\lambda}| = \left|\int_{\mathbb{M}}\widetilde{\chi}_{\lambda}d\mu\right| \le \|\widetilde{\chi}_{\lambda}\|_{L_1} \le Cq_{\Lambda}^{d}.$$

Hence,

$$\left|B_{\xi\eta}^{\Lambda}\right| \le Ch_{\Xi}^{-2}q_{\Lambda}^{d}\sum_{\lambda\in\Lambda}\left(1+\frac{\operatorname{dist}(\lambda,\xi)}{h_{\Xi}}\right)^{-2m}\left(1+\frac{\operatorname{dist}(\lambda,\eta)}{h_{\Xi}}\right)^{-2m}. \quad (4.1.2)$$

Let

$$\Lambda_{\xi} = \{\lambda\in\Lambda : \operatorname{dist}(\lambda,\xi) < \operatorname{dist}(\lambda,\eta)\}$$

and

$$\Lambda_{\eta} = \{\lambda\in\Lambda : \operatorname{dist}(\lambda,\eta) \le \operatorname{dist}(\lambda,\xi)\}.$$

Note that for $\lambda\in\Lambda_{\xi}$, $\operatorname{dist}(\lambda,\eta) \ge \frac{1}{2}\operatorname{dist}(\xi,\eta)$ and so

$$\begin{aligned}&\sum_{\lambda\in\Lambda_{\xi}}\left(1+\frac{\operatorname{dist}(\lambda,\xi)}{h_{\Xi}}\right)^{-2m}\left(1+\frac{\operatorname{dist}(\lambda,\eta)}{h_{\xi}}\right)^{-2m} \\ &\le C\left(1+\frac{\operatorname{dist}(\xi,\eta)}{h_{\xi}}\right)^{-2m}\sum_{\lambda\in\Lambda}\left(1+\frac{\operatorname{dist}(\lambda,\xi)}{h_{\Xi}}\right)^{-2m}.\end{aligned} \quad (4.1.3)$$

Note that if $p \in b(\lambda, q_\Lambda)$, then $\text{dist}(\lambda, \xi) \geq \frac{1}{2}\text{dist}(p, \xi)$, and hence

$$\left(1 + \frac{\text{dist}(\lambda, \xi)}{h_\Xi}\right)^{-2m} \leq C \left(1 + \frac{\text{dist}(p, \xi)}{h_\Xi}\right)^{-2m}.$$

Therefore,

$$\begin{aligned}\left(1 + \frac{\text{dist}(\lambda, \xi)}{h_\Xi}\right)^{-2m} &= \left(1 + \frac{\text{dist}(\lambda, \xi)}{h_\Xi}\right)^{-2m} \cdot q_\Lambda^{-d} \int_{b(\lambda, q_\Lambda)} d\mu \\ &\leq C q_\Lambda^{-d} \int_{b(\lambda, q_\Lambda)} \left(1 + \frac{\text{dist}(p, \xi)}{h_\Xi}\right)^{-2m} d\mu(p).\end{aligned}$$

Since the collection of balls of radius q_Λ centered at $\lambda \in \Lambda$ are disjoint, we have

$$\begin{aligned}&\sum_{\lambda \in \Lambda} \left(1 + \frac{\text{dist}(\lambda, \xi)}{h_\Xi}\right)^{-2m} \\ &\leq C q_\Lambda^{-d} \sum_{\lambda \in \Lambda} \int_{b(\lambda, q_\Lambda)} \left(1 + \frac{\text{dist}(p, \xi)}{h_\Xi}\right)^{-2m} d\mu(p) \qquad (4.1.4) \\ &\leq C q_\Lambda^{-d} \int_{\mathbb{M}} \left(1 + \frac{\text{dist}(p, \xi)}{h_\Xi}\right)^{-2m} d\mu(p) \leq C q_\Lambda^{-d} h_\Xi^d,\end{aligned}$$

since the integral in the last line of (4.1.4) is precisely the integral we showed was bounded by $C h_\Xi^d$ in the proof of Lemma 3.4.1. Putting (4.1.4) into (4.1.3) and then putting that into (4.1.2) yields the result. The result in the case that Φ_m provides an exponential energy estimate is proven in an almost identical way. □

Theorem 4.1.5. *Suppose Φ_m provides an energy estimate with parameter $m > d/2 + 1$, Ξ is a finite subset of $\mathbb{M}$ (Π-unisolvent if Φ_m is conditionally positive definite with respect to Π), and $\{\chi_\xi\}_{\xi \in \Xi}$ is the corresponding Lagrange basis. Then $b\chi_\xi\chi_\eta \in H^m \cap L_\infty$ and $a^\sharp(\nabla\chi_\xi, \nabla\chi_\eta) \in H^{m-1}$; more-*

over, there is a positive constant C such that, for h_Ξ sufficiently small,

$$\|b\chi_\xi\chi_\eta\|_{H^m} \le Ch_\Xi^{\frac{d}{2}-m} \tag{4.1.6}$$

and

$$\left\|a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right)\right\|_{H^{m-1}} \le Ch_\Xi^{\frac{d}{2}-m-1}. \tag{4.1.7}$$

Proof. By Lemma 2.7.5,

$$\|b\chi_\xi\chi_\eta\|_{H^m} \le C\left(\|b\|_{H^m}\|\chi_\xi\chi_\eta\|_{L_\infty} + \|b\|_{L_\infty}\|\chi_\xi\chi_\eta\|_{H^m}\right).$$

We know $\|b\|_{L_\infty} \le \|b\|_{H^m} = C$, and $\|\chi_\xi\chi_\eta\|_{L_\infty} \le \|\chi_\xi\|_{L_\infty}\|\chi_\eta\|_{L_\infty} \le C$. Another application of Lemma 2.7.5 then gives

$$\|b\chi_\xi\chi_\eta\|_{H^m} \le C\left(1 + \left(\|\chi_\xi\|_{H^m}\|\chi_\eta\|_{L_\infty} + \|\chi_\xi\|_{L_\infty}\|\chi_\eta\|_{H^m}\right)\right).$$

Again $\|\chi_\xi\|_{L_\infty} \le C$, and our bump estimate gives $\|\chi_\xi\|_{H^m} \le Cq_\xi^{\frac{d}{2}-m}$. This proves (4.1.6).

The proof of (4.1.7) is somewhat more involved. We start by covering $\mathbb{M}$ with finitely many coordinate patches $\Omega_1, \ldots, \Omega_K$, where

$$\Omega_k \subseteq b(q_k, r_{\mathbb{M}}/3)$$

for each k and $q_k \in \mathbb{M}$. Here K is a fixed constant guaranteed by the compactness of $\mathbb{M}$. We can use normal coordinates $\mathrm{Exp}_k = \mathrm{Exp}_{q_k}$ for each k, since $\mathrm{Exp}_k : U_k \to \Omega_k$ is bijective (here $U_k = \mathrm{Exp}_k^{-1}(\Omega_k)$). Let $\tau_1, \ldots, \tau_K$

be a partition of unity subordinate to $\Omega_1, \ldots, \Omega_K$. Then

$$\begin{aligned} \left\| a^\sharp \left(\nabla\chi_\xi, \nabla\chi_\eta \right) \right\|^2_{H^{m-1}} &\leq C \left\| a^\sharp \left(\nabla\chi_\xi, \nabla\chi_\eta \right) \right\|^2_{W_2^{m-1}} \\ &\leq C \sum_{k=1}^{K} \left\| \tau_k a^\sharp \left(\nabla\chi_\xi, \nabla\chi_\eta \right) \right\|^2_{W_2^{m-1}(\Omega_k)} . \end{aligned} \tag{4.1.8}$$

Using the metric equivalence from Lemma 2.6.1,

$$\left\| \tau_k a^\sharp \left(\nabla\chi_\xi, \nabla\chi_\eta \right) \right\|_{W_2^{m-1}(\Omega_k)} \leq C \left\| \tau_k \sum_{i,j} a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j} \right\|_{W_2^{m-1}(U_k)}, \tag{4.1.9}$$

where we have abused notation slightly by writing, for example, a^{ij} instead of $a^{ij} \circ \mathrm{Exp}_k^{-1}$. Thus by the triangle inequality,

$$\left\| \tau_k a^\sharp \left(\nabla\chi_\xi, \nabla\chi_\eta \right) \right\|_{W_2^{m-1}(\Omega_k)} \leq \sum_{i,j} \left\| \tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j} \right\|_{W_2^{m-1}(U_k)} . \tag{4.1.10}$$

Using the metric equivalence from Lemma 2.6.1 again,

$$\left\| \tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j} \right\|_{W_2^{m-1}(U_k)} \leq C \left\| \tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j} \right\|_{H^{m-1}}, \tag{4.1.11}$$

where by $\tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j}$ we mean the extension by zero of this function from Ω_k to all of $\mathbb{M}$. This is justified because the closure of the support of τ_k is contained in Ω_k. Let σ be a smooth cutoff function with $\sigma = 1$ on the support of τ_k and $\sigma = 0$ on $\mathbb{M} \setminus \Omega_k$. We now apply Lemma 2.7.5 twice to obtain

$$
\begin{aligned}
&\left\|\tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i} \frac{\partial \chi_\eta}{\partial x^j}\right\|_{H^{m-1}} \\
&\leq C\left(\left(\left\|\tau_k a^{ij}\right\|_{H^{m-1}} \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{L_\infty}\right.\right. \\
&\qquad\qquad \left. + \left\|\tau_k a^{ij}\right\|_{L_\infty} \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}\right) \left\|\frac{\partial \chi_\eta}{\partial x^j}\right\|_{L_\infty} \\
&\qquad\qquad\qquad \left. + \left\|\tau_k a^{ij} \frac{\sigma \partial \chi_\xi}{\partial x^i}\right\|_{L_\infty} \left\|\sigma \frac{\partial \chi_\eta}{\partial x^j}\right\|_{H^{m-1}}\right).
\end{aligned}
\tag{4.1.12}
$$

The H^{m-1} and L_∞ norms of $\tau_k a^{ij}$ can be absorbed into the inner constant, since they depend only on a and the choice of partition of unity. Similarly,

$$
\left\|\tau_k a^{ij} \frac{\partial \chi_\xi}{\partial x^i}\right\|_{L_\infty} \leq \left\|\tau_k a^{ij}\right\|_{L_\infty} \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{L_\infty} \leq C h_\Xi^{-1}
$$

by Lemma 3.2.7, since $\left\|\frac{\partial \chi_\xi}{\partial x^i}\right\|_{L_\infty} \leq \left\||\nabla \chi_\xi|\right\|_{L_\infty}$. (Because $\left|\frac{\partial \chi_\xi}{\partial x^i}\right|$ is just one of the directional derivatives that $|\nabla \chi_\xi|$ is the maximum of.) The last norms in (4.1.8) to handle are the H^{m-1} norms of the partial derivatives of the Lagrange functions.

This will require yet another pass through Euclidean space using the metric equivalence from Lemma 2.6.1 and the norm equivalence from Lemma 2.7.3. First, the norm equivalence gives

$$
\begin{aligned}
\left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}^2 &\leq C \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{W_2^{m-1}(\mathbb{M})}^2 \\
&= C \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{W_2^{m-1}(\Omega_k)}^2 .
\end{aligned}
$$

The metric equivalence now gives

$$\left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}^2 \le C \left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{W_2^{m-1}(U_k)}^2,$$

where we have again abused notation slightly by writing $\frac{\partial \chi_\xi}{\partial x^i}$ instead of $\frac{\partial \chi_\xi}{\partial x^i} \circ \operatorname{Exp}_k^{-1}$ on the left and σ instead of $\sigma \circ \operatorname{Exp}_k$. By the definition of $W_2^{m-1}(U_k)$ and Leibniz's rule,

$$\begin{aligned}\left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}^2 &\le C \sum_{|\alpha| \le m-1} \int_{U_k} \left| D^\alpha \left(\sigma(x) \frac{\partial \chi_\xi}{\partial x^i}(x) \right) \right|^2 dx \\ &= C \sum_{|\alpha| \le m-1} \int_{U_k} \left| \sum_{\beta \le \alpha} D^\beta \sigma(x) D^{\alpha-\beta} \frac{\partial \chi_\xi}{\partial x^i}(x) \right|^2 dx.\end{aligned}$$

Now, all the derivatives of σ are bounded, as is the number of times each γ with $|\gamma| \le m-1$ appears. Absorbing these bounds into the constant, we have

$$\left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}^2 \le C \sum_{|\gamma| \le m-1} \int_{U_k} \left| D^\gamma \frac{\partial \chi_\xi}{\partial x^i} \right|^2 dx.$$

For each such γ, $D^\gamma \frac{\partial \chi_\xi}{\partial x_i} = D^\delta \chi_\xi$ for some δ with $|\delta| \le m$. Hence

$$\begin{aligned}\left\|\sigma \frac{\partial \chi_\xi}{\partial x^i}\right\|_{H^{m-1}}^2 &\le C \sum_{|\delta| \le m} \int_{U_k} \left| D^\delta \chi_\xi \right|^2 dx = C \left\|\chi_\xi\right\|_{W_2^m(U_k)}^2 \\ &\le C \left\|\chi_\xi\right\|_{W_2^m(\mathbb{M})}^2 \le C \left\|\chi_\xi\right\|_{H^m}^2 \le C q_\Xi^{d-2m},\end{aligned} \tag{4.1.13}$$

where we have used the norm equivalence in the second-to-last inequality and our bump estimate in the last. The result now follows by successively plugging back in to (4.1.12), (4.1.11), (4.1.10), (4.1.9), and (4.1.8). □

Lemma 4.1.14. *Suppose Φ_m provides an energy estimate with parameter $m > \frac{d}{2} + 1$, and let Ξ and Λ be finite sets of centers (both Π_L-unisolvent if*

Φ_m is conditionally positive definite with respect to Π_L). Assume $\#\Xi < \#\Lambda$ and $h_\Lambda < h_\Xi$. Let B_Ξ and B_Ξ^Λ be the stiffness matrix and quadratized stiffness matrix, respectively, where Ξ is the set of centers use for Galerkin approximation and Λ is the set of centers used for quadrature. There is a constant C such that, for h_Ξ sufficiently small, the uniform estimate

$$\left|B_{\xi\eta} - B_{\xi\eta}^\Lambda\right| \le C\left(\frac{h_\Lambda}{h_\Xi}\right)^m h_\Xi^{\frac{d}{2}-1} h_\Lambda^{-1},$$

holds for all $\xi, \eta \in \Xi$.

Proof. Our quadrature rule, Theorem 3.1.1, along with Theorem 4.1.5 gives

$$\begin{aligned}\left|\int_{\mathbb{M}} b\chi_\xi\chi_\eta d\mu - Q^\Lambda\left(b\chi_\xi\chi_\eta\right)\right| &\le Ch_\Lambda^m \left\|b\chi_\xi\chi_\eta\right\|_{H^m} \\ &\le Ch_\Lambda^m h_\Xi^{\frac{d}{2}-m}.\end{aligned}$$

Similarly,

$$\begin{aligned}&\left|\int_{\mathbb{M}} a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right) d\mu - Q^\Lambda\left(a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right)\right)\right| \\ &\le Ch_\Lambda^{m-1}\left\|a^\sharp\left(\nabla\chi_\xi, \nabla\chi_\eta\right)\right\|_{H^{m-1}} \\ &\le Ch_\Lambda^{m-1} h_\Xi^{\frac{d}{2}-m-1},\end{aligned}$$

and the result follows by the triangle inequality. □

Corollary 4.1.15. *Adopt the notation and assumptions from Lemma 4.1.14. There is a constant C such that, for h_Ξ sufficiently small,*

$$\left\|B_\Xi - B_\Xi^\Lambda\right\|_1 \le C\left(\frac{h_\Lambda}{h_\Xi}\right)^m h_\Xi^{-\frac{d}{2}-1} h_\Lambda^{-1}.$$

Proof. We have $\left\|B_\Xi - B_\Xi^\Lambda\right\|_1 = \max_{\xi\in\Xi}\sum_{\eta\in\Xi}\left|B_{\xi\eta} - B_{\xi\eta}^\Lambda\right|$. By the lemma, for fixed $\xi \in \Xi$ we have

$$\sum_{\eta\in\Xi} \left|B_{\xi\eta} - B^{\Lambda}_{\xi\eta}\right| \le (\#\Xi)\cdot C\left(\frac{h_\Lambda}{h_\Xi}\right)^m h_\Xi^{\frac{d}{2}-1} h_\Lambda^{-1},$$

and $\#\Xi \sim h_\Xi^{-d}$ yields the result. □

The uniform estimate from Lemma 4.1.14 allows us to prove Corollary 4.1.15; however, we are not taking advantage of our energy estimates. To do so, we fix $\xi \in \Xi$ and break the sum in the proof into an "inner sum" and an "outer sum". The uniform estimate will be used on the inner sum, whereas the bounds on the entries of the stiffness and quadratized stiffness matrices from Lemmas 3.4.1 and 4.1.1 will be used on the outer sum.

Lemma 4.1.16. *Adopt the notation and assumptions from Lemma 4.1.14. Suppose Φ provides an algebraic energy estimate as in (3.2.3). Assume in addition that*

$$h_\Lambda \ge \operatorname{diam}(\mathbb{M})^{-2-\frac{2}{m-1}} h_\Xi^{3-\frac{d-8}{2m-2}}. \tag{4.1.17}$$

There is a positive constant C such that, for h_Ξ sufficiently small,

$$\left\|B_\Xi - B^\Lambda_\Xi\right\|_1 \le C\left(\frac{h_\Lambda}{h_\Xi}\right)^{m-\frac{d}{2}-1} h_\Xi^{\frac{d}{2}-2}.$$

Proof. We want to bound $\left\|B_\Xi - B^\Lambda_\Xi\right\|_1 = \max_{\xi\in\Xi}\sum_{\eta\in\Xi}\left|B_{\xi\eta} - B^\Lambda_{\xi\eta}\right|$. Fixing $\xi \in \Xi$ and $r_0 > 0$, we split the sum into an inner sum and an outer sum:

$$\sum_{\eta\in\Xi}\left|B_{\xi\eta} - B^\Lambda_{\xi\eta}\right| = \sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)}\left|B_{\xi\eta} - B^\Lambda_{\xi\eta}\right| + \sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)^c}\left|B_{\xi\eta} - B^\Lambda_{\xi\eta}\right|.$$

Let us consider the outer sum first. Divide $\mathbb{M}$ into annuli A_n with outer radius nh_Ξ and inner radius $(n-1)h_\Xi$, $n = 1, \dots, n_{\max}$, and let $n_0 = \lfloor \frac{r_0}{h_\Xi} \rfloor$.

By Lemmas 3.4.1 and 4.1.1,

$$\begin{aligned}
\sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)^c}\left|B_{\xi\eta}-B_{\xi\eta}^{\Lambda}\right| &\le Ch_\Xi^{d-2}\sum_{n=n_0}^{n_{\max}}\sum_{\eta\in\Xi\cap A_n}\left(1+\frac{\operatorname{dist}(\xi,\eta)}{h_\Xi}\right)^{-2m}\\
&\le Ch_\Xi^{d-2}\sum_{n=n_0}^{n_{\max}}\left(\#\Xi\cap A_n\right)\left(1+\frac{(n-1)h_\Xi}{h_\Xi}\right)^{-2m}\\
&= Ch_\Xi^{d-2}\sum_{n=n_0}^{n_{\max}}\left(\#\Xi\cap A_n\right)n^{-2m}.
\end{aligned}$$

Now, $\operatorname{Vol}(A_n)\sim n^{d-1}h_\Xi^d$, and hence $\#\Xi\cap A_n\sim\frac{n^{d-1}h_\Xi^d}{q_\Xi^d}=\rho^d n^{d-1}$. Therefore

$$\begin{aligned}
\sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)^c}\left|B_{\xi\eta}-B_{\xi\eta}^{\Lambda}\right| &\le Ch_\Xi^{d-2}\sum_{n=n_0}^{n_{\max}}n^{d-2m-1}\\
&\le Ch_\Xi^{d-2}\sum_{n=n_0}^{\infty}n^{d-2m-1}.
\end{aligned}$$

An integral comparison gives

$$\begin{aligned}
\sum_{n=n_0}^{\infty}n^{d-2m-1} &\le n_0^{d-2m-1}\left(1+\frac{n_0}{2m-d}\right)\\
&\le h_\Xi^{2m-d+1}r_0^{d-2m-1}\left(1+\frac{r_0}{2m-d}\right).
\end{aligned}$$

Hence we obtain

$$\sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)^c}\left|B_{\xi\eta}-B_{\xi\eta}^{\Lambda}\right|\le Ch_\Xi^{2m-1}r_0^{d-2m-1}\left(1+\frac{r_0}{2m-d}\right). \tag{4.1.18}$$

Let us now examine the inner sum. The set of centers $\Xi\cap\mathbf{b}(\xi,r_0)$ has cardinality bounded by

$$\frac{\operatorname{Vol}\left(\mathbf{b}\left(\xi,r_0\right)\right)}{\operatorname{Vol}\left(\mathbf{b}\left(\xi,q_\Xi\right)\right)}\sim\frac{r_0^d}{q_\Xi^d}=\rho^d r_0^d h_\Xi^{-d}.$$

The uniform estimate from Lemma 4.1.14, then, gives

$$\sum_{\eta\in\Xi\cap\mathbf{b}(\xi,r_0)} \left|B_{\xi\eta}-B^{\Lambda}_{\xi\eta}\right| \le C h_{\Xi}^{-\frac{d}{2}-m-1} h_{\Lambda}^{m-1} r_0^{d}. \tag{4.1.19}$$

Assume for now that $r_0 \le \operatorname{diam}(\mathbb{M})$. Then (4.1.18) becomes

$$\begin{aligned}\sum_{\eta\in\mathbf{b}(\xi,r_0)^{\mathrm{c}}} \left|B_{\xi\eta}-B^{\Lambda}_{\xi\eta}\right| &\le C h_{\Xi}^{2m-1} r_0^{d-2m-1}\left(1+\frac{\operatorname{diam}(\mathbb{M})}{2m-d}\right)\\ &= C h_{\Xi}^{2m-1} r_0^{d-2m-1}.\end{aligned} \tag{4.1.20}$$

Setting the results from (4.1.19) and (4.1.20) equal and solving for r_0 gives

$$r_0 = h_{\Xi}^{\frac{6m+d}{4m+2}} h_{\Lambda}^{-\frac{m-1}{2m+1}} = h_{\Xi}^{\frac{3}{2}+\frac{d-3}{4m+2}} h_{\Lambda}^{-\frac{1}{2}+\frac{3}{4m+2}}.$$

The condition on h_Λ, (4.1.17), ensures $r_0 \le \operatorname{diam}(\mathbb{M})$. Therefore (4.1.20) holds for this r_0, and plugging it into either (4.1.19) or (4.1.20) yields the result. □

Lemma 4.1.21. *Adopt the notation and assumptions from Lemma 4.1.14. Suppose Φ provides an exponential energy estimate as in (3.2.4). Assume in addition that*

$$\nu^{\frac{1}{m-1}} h_{\Xi}^{1+\frac{d-4}{2m-2}} e^{-\operatorname{diam}(\mathbb{M})/\nu h_{\Xi}(m-1)} \le h_\Lambda < \left(\frac{\nu}{e}\right) h_{\Xi}^{1+\frac{d-4}{2m-2}}. \tag{4.1.22}$$

There is a positive constant C such that, for h_Ξ sufficiently small,

$$\left\|B_{\Xi}-B^{\Lambda}_{\Xi}\right\|_1 \le C\left(\frac{h_\Lambda}{h_\Xi}\right)^{m-1} h_{\Xi}^{\frac{d}{2}-2} h_{\Lambda}^{-\epsilon},$$

where $\epsilon = \frac{d\log|\log(h_\Lambda)|}{|\log(h_\Lambda)|}$.

Proof. We again have $\left\|B_{\Xi}-B^{\Lambda}_{\Xi}\right\|_1 = \max_{\xi\in\Xi}\sum_{\eta\in\Xi}\left|B^{\Lambda}_{\xi\eta}-B^{\Lambda}_{\xi\eta}\right|$, and for a fixed

$\xi \in \Xi$ we split the sum into an inner sum and an outer sum as in the proof of Lemma 4.1.16. Handling the outer sum first, we have by Lemmas 3.4.1 and 4.1.1 that

$$\begin{aligned}
\sum_{\eta \in \Xi \cap \mathbf{b}(\xi, r_0)^c} \left| B_{\xi\eta} - B^{\Lambda}_{\xi\eta} \right| &\leq C h_{\Xi}^{d-2} \sum_{n=n_0}^{n_{\max}} \sum_{\eta \in \Xi \cap A_n} \exp\left(-\nu \frac{\operatorname{dist}(\xi, \eta)}{h_{\Xi}} \right). \\
&\leq C h_{\Xi}^{d-2} \sum_{n=n_0}^{n_{\max}} \left(\#(\Xi \cap A_n) \right) e^{-\nu(n-1)} \\
&\leq C h_{\Xi}^{d-2} \sum_{n=n_0}^{\infty} n^{d-1} e^{-\nu(n-1)}.
\end{aligned}$$

An integral comparison now gives

$$\begin{aligned}
\sum_{n=n_0}^{\infty} n^{d-1} e^{-\nu(n-1)} &\leq e^{-\nu(n_0-1)} \left(n_0^{d-1} + \sum_{i=0}^{d-1} (-1)^{d-i} \frac{(d-1)!}{i!(-\nu)^{d-i}} n_0^i \right) \\
&\leq C(d-1) n_0^{d-1} e^{-\nu n_0} \\
&\leq C' h_{\Xi}^{1-d} r_0^{d-1} e^{-\nu r_0 / h_{\Xi}}.
\end{aligned}$$

where we have taken

$$C = e^{\nu} \max\left\{ 1, \max\left\{ \left| \frac{(d-1)!}{i!(-\nu)^{d-i}} \right| \, : \, i \in \{0, \ldots, d-1\} \right\} \right\}.$$

We thus arrive at

$$\sum_{\eta \in \mathbf{b}(\xi, r_0)^c} \left| B_{\xi\eta} - B^{\Lambda}_{\xi\eta} \right| \leq C h_{\Xi}^{-1} r_0^{d-1} e^{-\nu r_0 / h_{\Xi}}. \tag{4.1.23}$$

The bound in (4.1.19) remains the same. Our task now is to find the value of r_0 that makes (4.1.19) and (4.1.23) equal. This requires solving an equation of the form $a r_0 = e^{b r_0}$, where $a = h_{\Xi}^{-\frac{d}{2}-m} h_{\Lambda}^{m-1}$ and $b = -\nu h_{\Xi}^{-1}$. Since $\frac{b}{a} = -\nu h_{\Xi}^{m+\frac{d}{2}-1} h_{\Lambda}^{1-m} < 0$, there is a unique solution; namely $r_0 = -\frac{1}{b} W\left(-\frac{b}{a}\right)$, where W is the principal branch of the Lambert W function. Hence,

$$r_0 = \nu^{-1} h_\Xi W\left(\nu h_\Xi^{m+\frac{d}{2}-1} h_\Lambda^{1-m}\right). \tag{4.1.24}$$

The lower bound on h_Λ in (4.1.22) ensures that $r_0 \leq \mathrm{diam}(\mathbb{M})$. Plugging (4.1.24) into either (4.1.19) or (4.1.23) yields

$$\left\|B_\Xi - B_\Xi^\Lambda\right\|_1 \leq C h_\Xi^{\frac{d}{2}-m-1} h_\Lambda^{m-1} W\left(\nu h_\Xi^{m+\frac{d}{2}-1} h_\Lambda^{1-m}\right)^d.$$

The upper bound on h_Λ in (4.1.22) ensures $\nu h_\Xi^{m+\frac{d}{2}-1} h_\Lambda^{1-m} > e$, and by [19, Theorem 2.1] $W(t) < \log(t)$ for $t > e$. Hence the result now follows from

$$\begin{aligned}
&\left|\log\left(\nu h_\Xi^{m+\frac{d}{2}-1} h_\Lambda^{1-m}\right)\right| \\
&= \left|\log(\nu) + \left(m + \frac{d}{2} - 1\right)\log(h_\Xi) - (m-1)\log(h_\Lambda)\right| \\
&\leq C\left|\log(h_\Lambda)\right|
\end{aligned}$$

□

Lemma 4.1.25. *Adopt the notation and assumptions from Lemma 4.1.14. Let* $C_0 = \max\{C_1, C_2, C_3, 1\}$*, where* C_1*,* C_2*, and* C_3 *are the constants from Lemmas 3.4.4, 4.1.16, and 4.1.21, respectively. Assume* h_Λ *is small enough that*

$$h_\Lambda \leq C_0^{\frac{4}{2+d-2m}} h_\Xi^{1+\frac{4d+4}{2m-d-2}} 2^{\frac{2}{2+d-2m}}.$$

There is a constant C *such that, for* h_Ξ *sufficiently small,*

$$\left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2 \leq C h_\Xi^{-d}.$$

Remark 4.1.26. Before we prove the Lemma, we explain the upper bound on h_Λ. In [22], the authors make a the assumption that the quadrature weights

$\{w_\zeta\}_{\zeta\in\Lambda}$ ($\{w_y\}_{y\in Y}$ in their notation) are not only positive, but bounded below by Ch_Λ^d (Ch_Y^2 in their notation) - see [22, Assumption 4.15]. This is a conjecture for quadrature on quasi-uniform sets of centers supported by some amount of empirical data, but has yet to be proven, even on $\mathbb{S}^2$. If that assumption is made, then the result in Lemma 4.1.25 can be achieved without placing the restrictive upper bound on h_Λ - the proof is analogous to the proof found in [22, Theorem 7.7]. We will revisit these ideas in Remark 4.2.9, and give the analogous proof in Lemma 4.2.12.

Remark 4.1.27. We note that the 2 in the upper bound for h_Λ was chosen for simplicity. It can be replaced by $1+\epsilon$ for any $\epsilon > 0$. The idea is to ensure the convergence of the Neumann series that arises in the proof.

Proof. Note that

$$B_\Xi^\Lambda = B_\Xi\left(I - B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right),$$

and hence

$$\left(B_\Xi^\Lambda\right)^{-1} = \left(I - B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right)^{-1} B_\Xi^{-1}.$$

The matrices B_Ξ^{-1} and $B_\Xi - B_\Xi^\Lambda$ are self-adjoint, and hence $\left\|B_\Xi^{-1}\right\|_2 \le \left\|B_\Xi^{-1}\right\|_1$ and $\left\|B_\Xi - B_\Xi^\Lambda\right\|_2 \le \left\|B_\Xi - B_\Xi^\Lambda\right\|_1$. By Lemmas 3.4.4 and 4.1.16, then,

$$\begin{aligned}\left\|B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right\|_2 &\le \left\|B_\Xi^{-1}\right\|_2 \left\|B_\Xi - B_\Xi^\Lambda\right\|_2 \\ &\le C_0^2 h_\Xi^{-m-\frac{3d}{2}-1} h_\Lambda^{m-\frac{d}{2}-1}.\end{aligned}$$

The condition on h_Λ ensures that

$$\left\|B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right\|_2 \le \frac{1}{2},$$

and this in turn ensures the convergence of the Neumann series

$$\left(I - B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right)^{-1} = \sum_{n=0}^{\infty}\left(B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right)^n.$$

Therefore

$$\left\|\left(I - B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right)^{-1}\right\|_2 \le \sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^n = 2,$$

and hence

$$\left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2 \le \left\|\left(I - B_\Xi^{-1}\left(B_\Xi - B_\Xi^\Lambda\right)\right)^{-1}\right\|_2 \left\|B_\Xi^{-1}\right\|_2 \le C h_\Xi^{-d}$$

by another application of Lemma 3.4.4, since $\left\|B_\Xi^{-1}\right\|_2 \le \left\|B_\Xi^{-1}\right\|_1$ as well.

□

4.2 Error Estimates

We are now ready to prove error estimates for our quadratized Galerkin approximation to $\mathcal{L}u = f$. To recapitulate, the Galerkin approximation is

$$u_\Xi = \sum_{\xi\in\Xi} \gamma_\xi \chi_\xi,$$

whose coefficients $\gamma = \{\gamma_\xi\}_{\xi\in\Xi}$ are obtained by solving the linear system $B_\Xi\gamma = \omega$, where B_Ξ is the stiffness matrix and $\omega = \{\omega_\xi\}_{\xi\in\Xi}$ has entries

$$\omega_\xi = \int_{\mathbb{M}} \chi_\xi f \, d\mu.$$

The quadratized Galerkin approximation,

$$u_\Xi^\Lambda = \sum_{\xi\in\Xi} \gamma_\xi^\Lambda \chi_\xi,$$

has coefficients $\gamma^\Lambda = \left\{\gamma^\Lambda_\xi\right\}_{\xi\in\Xi}$ obtained by solving the system $B^\Lambda_\Xi \gamma^\Lambda = \omega^\Lambda$, where $\omega^\Lambda = \left\{\omega^\Lambda_\xi\right\}_{\xi\in\Xi}$ is the vector with entries $\omega^\Lambda_\xi = Q^\Lambda\left(\chi_\xi f\right)$.

In Theorem 4.2.1, we estimate the error between the Galerkin approximation and the quadratized Galerkin approximation. Corollary 4.2.7 then gives the full error estimate between the weak solution to $\mathcal{L}u = f$ and the quadratized Galerkin approximation. It is simply the result of an application of the triangle inequality.

Theorem 4.2.1. *Suppose Φ_m provides an energy estimate with parameter $m > \frac{d}{2} + 1$, and let Ξ and Λ be finite subsets of $\mathbb{M}$ (both Π-unisolvent if Φ_m is conditionally positive definite with respect to Π). Assume $\#\Xi < \#\Lambda$, and that h_Λ satisfies the upper bound from Lemma 4.1.25. Let $f \in H^s$, where $s = m - 1$ if $\frac{d}{2} + 1 < m \leq \frac{d}{2} + 2$ and $s = m - 2$ if $m > \frac{d}{2} + 2$. Let u^Λ_Ξ be the quadratized Galerkin approximation to $\mathcal{L}u = f$, using Ξ for Galerkin approximation and Λ for quadrature.*

Suppose Φ_m provides an algebraic energy estimate as in (3.2.3). Assume h_Λ also satisfies the bound from Lemma 4.1.16. There is a positive constant C such that

$$\left\|u_\Xi - u^\Lambda_\Xi\right\|_{L_2} \leq C \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-\frac{d}{2}-1} h_\Xi^{-\frac{d}{2}-2} \left\|f\right\|_{H^s}.$$

Now suppose Φ_m provides an exponential energy estimate as in (3.2.4). Assume h_Λ also satisfies the bounds from Lemma 4.1.21. There is a positive constant C such that

$$\left\|u_\Xi - u^\Lambda_\Xi\right\|_{L_2} \leq C \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-1} h_\Xi^{-\frac{d}{2}-2} h_\Lambda^{-\epsilon} \left\|f\right\|_{H^s},$$

where $\epsilon = \frac{d\,\log|\log(h_\Lambda)|}{|\log(h_\Lambda)|}$.

Remark 4.2.2. Before we prove the result, we explain why there are different cases for s. If $\frac{d}{2} + 1 < m \leq \frac{d}{2} + 2$, then we have a problem if we take $f \in H^{m-2}$, because $m - 2 \leq \frac{d}{2}$, and the Sobolev embedding theorem doesn't guarantee $H^{m-2} \subset C(\mathbb{M})$. Hence, we don't know that $f \in H^{m-2}$ is a continuous function, or even that it is defined everywhere. In that case we don't know that we can sample f on Λ, which is necessary for quadrature. That is why we insist on $f \in H^{m-1}$ in that case, and not H^{m-2}.

Proof. By the L_2-stability of the Lagrange basis,

$$\begin{aligned}\left\|u_\Xi - u_\Xi^\Lambda\right\|_{L_2} &= \left\|\sum_{\xi\in\Xi} \left(\gamma_\xi - \gamma_\xi^\Lambda\right) \chi_\xi\right\|_{L_2} \\ &\leq C h_\Xi^{d/2} \left\|\gamma - \gamma^\Lambda\right\|_{\ell_2(\Xi)} \\ &= C h_\Xi^{d/2} \left\|B_\Xi^{-1}\omega - \left(B_\Xi^\Lambda\right)^{-1} \omega^\Lambda\right\|_{\ell_2(\Xi)}.\end{aligned}$$

Adding and subtracting $\left(B_\Xi^\Lambda\right)^{-1} \omega$ inside the norm and applying the triangle inequality gives

$$\begin{aligned}\left\|u_\Xi - u_\Xi^\Lambda\right\|_{L_2} \leq C h_\Xi^{d/2} \Bigg(&\left\|\left(B_\Xi^{-1} - \left(B_\Xi^\Lambda\right)^{-1}\right) \omega\right\|_{\ell_2(\Xi)} \\ &+ \left\|\left(B_\Xi^\Lambda\right)^{-1} \left(\omega - \omega^\Lambda\right)\right\|_{\ell_2(\Xi)}\Bigg).\end{aligned}$$

Noting that $B_\Xi^{-1} - \left(B_\Xi^\Lambda\right)^{-1} = \left(B_\Xi^\Lambda\right)^{-1} \left(B_\Xi^\Lambda - B_\Xi\right) B_\Xi^{-1}$, we have

$$\begin{aligned}\left\|u_\Xi - u_\Xi^\Lambda\right\|_{L_2} \leq C h_\Xi^{d/2} \Bigg(&\left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2 \left\|B_\Xi^\Lambda - B_\Xi\right\|_2 \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} \\ &+ \left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2 \left\|\omega - \omega^\Lambda\right\|_{\ell_2(\Xi)}\Bigg).\end{aligned}$$

Factoring out $\left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2$ and applying Lemma 4.1.25 gives

$$\left\|u_\Xi - u_\Xi^\Lambda\right\|_{L_2} \leq C h_\Xi^{-d/2} \left(\left\|B_\Xi^\Lambda - B_\Xi\right\|_2 \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} + \left\|\omega - \omega^\Lambda\right\|_{\ell_2(\Xi)} \right). \tag{4.2.3}$$

We bound the quantities in parentheses in (4.2.3) separately. For the first, we begin by using the L_2 stability of the Lagrange basis again to obtain

$$\begin{aligned} \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} = \|\gamma\|_{\ell_2(\Xi)} \leq C h_\Xi^{-d/2} \left\| \sum_{\xi\in\Xi} \gamma_\xi \chi_\xi \right\|_{L_2} &= C h_\Xi^{-d/2} \|u_\Xi\|_{L_2} \\ &\leq C h_\Xi^{-d/2} \left(\|u - u_\Xi\|_{L_2} + \|u\|_{L_2} \right), \end{aligned}$$

where u is the weak solution to $\mathcal{L}u = f$. By regularity $\|u\|_{L_2} \leq \|u\|_{H_2} \leq C\|f\|_{L_2} \leq C\|f\|_{H^s}$, and by Theorem 3.3.6, $\|u - u_\Xi\|_{L_2} \leq C h_\Xi^{m-1} \|f\|_{H^{m-2}} \leq C h_\Xi^{m-1}\|f\|_{H^s}$. Hence,

$$\left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} \leq C h_\Xi^{-d/2} \left(h_\Xi^{m-1} + 1 \right) \|f\|_{H^s} \leq C h_\Xi^{-d/2} \|f\|_{H^s}.$$

In the case of an algebraic energy estimate, we apply Lemma 4.1.16, keeping in mind that $\left\|B_\Xi - B_\Xi^\Lambda\right\|_2 \leq \left\|B_\Xi - B_\Xi^\Lambda\right\|_1$, to obtain

$$\left\|B_\Xi^\Lambda - B_\Xi\right\|_2 \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} \leq C \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-\frac{d}{2}-1} h_\Xi^{-2} \|f\|_{H^s}. \tag{4.2.4}$$

In the case of an exponential energy estimate, Lemma 4.1.21 gives

$$\left\|B_\Xi^\Lambda - B_\Xi\right\|_2 \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} \leq C \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-1} h_\Xi^{-2} h_\Lambda^{-\epsilon} \|f\|_{H^s}. \tag{4.2.5}$$

For the second quantity in the parentheses in (4.2.3), we begin with our quadrature rule and Lemma 2.7.5:

$$\left|\omega_\xi - \omega_\xi^\Lambda\right| \le C h_\Lambda^s \left(\|f\|_{H^s} \|\chi_\xi\|_{L_\infty} + \|f\|_{L_\infty} \|\chi_\xi\|_{H^s} \right).$$

Now, $\|f\|_{L_\infty} \le C\|f\|_{H^s}$, $\|\chi_\xi\|_{L_\infty} \le C$, and the Bernstein inequality from [14, Theorem 3] gives $\|\chi_\xi\|_{H^s} \le C h_\Xi^{\frac{d}{2}-s}$. Thus,

$$\left|\omega_\xi - \omega_\xi^\Lambda\right| \le C h_\Lambda^s \left(1 + h_\Xi^{\frac{d}{2}-s}\right) \|f\|_{H^s} \le C \left(\frac{h_\Lambda}{h_\Xi}\right)^s h_\Xi^{d/2}.$$

Taking the $\ell_2(\Xi)$-norm of these,

$$\left\|\omega - \omega^\Lambda\right\|_{\ell_2(\Xi)} \le (\#\Xi)^{1/2} \cdot C \left(\frac{h_\Lambda}{h_\Xi}\right)^s h_\Xi^{d/2} \|f\|_{H^s} \le C \left(\frac{h_\Lambda}{h_\Xi}\right)^s \|f\|_{H^s}. \tag{4.2.6}$$

In the algebraic case, (4.2.6) is clearly controlled by (4.2.4) for $s = m-1$ and $s = m-2$. If $s = m-1$, then (4.2.6) is clearly controlled by (4.2.5) as well. For the case $s = m-2$, we observe that $h_\Lambda/h_\Xi \le h_\Xi^{-2} h_\Lambda^{-\epsilon}$, since surely $h_\Lambda^{1+\epsilon} \le h_\Xi^{-1}$. Hence (4.2.6) is controlled by (4.2.5) in this case as well. The results now follow by putting (4.2.6) along with either (4.2.4) or (4.2.5) into (4.2.3).

□

Corollary 4.2.7. *Adopt the notation and assumptions from Theorem 4.2.1. Suppose Φ provides an algebraic energy estimate, and assume the bound on h_Λ from Lemma 4.1.16 holds. There is a positive constant C such that, for h_Ξ sufficiently small,*

$$\left\|u - u_\Xi^\Lambda\right\|_{L_2} \le C \left(h_\Xi^{m-1} + \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-\frac{d}{2}-1} h_\Xi^{-\frac{d}{2}-2} \right) \|f\|_{H^s}.$$

Now suppose Φ provides an exponential energy estimate, and assume *the bounds on h_Λ from Lemma 4.1.21 hold. There is a positive constant C such that, for h_Ξ sufficiently small,*

$$\left\|u - u_\Xi^\Lambda\right\|_{L_2} \leq C \left(h_\Xi^{m-1} + \left(\frac{h_\Lambda}{h_\Xi}\right)^{m-1} h_\Xi^{-\frac{d}{2}-2} h_\Lambda^{-\epsilon} \right) \|f\|_{H^s},$$

where $\epsilon = \frac{d\ log|log(h_\Lambda)|}{|log(h_\Lambda)|}$.

Remark 4.2.8. We come now to one of our main points: we can get the same result with quadrature as we can without by choosing h_Λ appropriately. Let us put aside the upper bound on h_Λ from Lemma 4.1.25 for a moment. We take as an ansatz that $h_\Lambda = h_\Xi^p$, and call p the *oversampling exponent.* The *optimal oversampling exponent* p^* is the smallest value of p that ensures that the quadrature error is on par with the Galerkin error. It is obtained by equating the quantities in the parentheses in Corollary 4.2.7 and solving for h_Λ. In the case of an algebraic energy estimate,

$$p^* = 2 + \frac{2d+4}{2m-d-2}.$$

For an exponential energy estimate,

$$p^* = 2 + \frac{d+4+4\epsilon}{2m-2-2\epsilon}.$$

Remark 4.2.9. Let us examine again the restriction on h_Λ in Lemma 4.1.25: $h_\Lambda \leq C_1 h_\Xi^{1+\frac{4d+4}{2m-d-2}}$, with

$$C_1 = C_0^{\frac{4}{2+d-2m}} 2^{\frac{2}{2+d-2m}}.$$

We know that C_0 depends only on Ξ via the mesh ratio, but we don't know what it is exactly. It is possible that this restriction is already stronger than the one imposed by the optimal oversampling exponent. If, however,

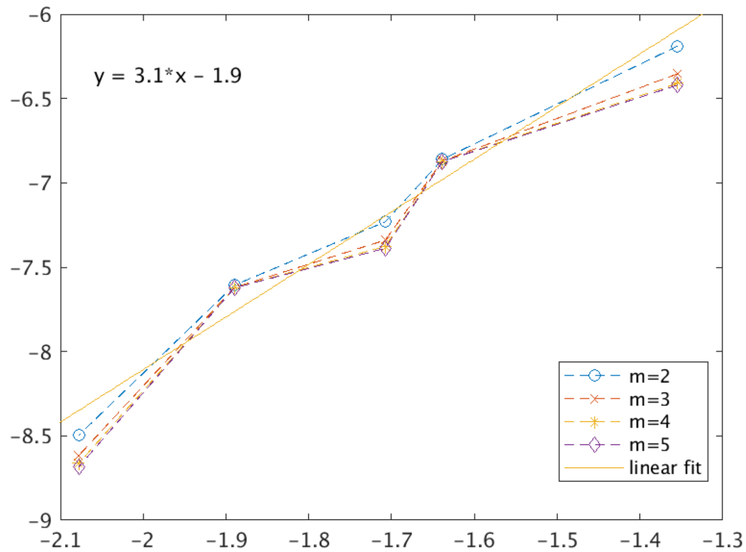
h_Λ is less than C_1 (the lemma does say "for h_Λ sufficiently small", after all) then the restriction becomes $h_\Lambda \le h_\Xi^{p_*}$, with the *baseline oversampling exponent*

$$p_* = 1 + \frac{4d+4}{2m-d-2}.$$

Noting that in this case $p_* \le p^*$, we now get a range for the oversampling exponent; namely, $p \in [p_*, p^*]$. If $p < p_*$, then Lemma 4.1.25, and hence Corollary 4.2.7, doesn't hold. Taking $p > p^*$ doesn't fundamentally improve the error estimate, since the Galerkin error stays the same. We note that with p in the range $[p_*, p^*]$, all of the bounds on h_Λ in Lemmas 4.1.16, 4.1.21, and 4.1.25 hold if $h_\Lambda = h_\Xi^p$, provided h_Ξ is sufficiently small.

Remark 4.2.10. With an additional assumption we can relax the bound on h_Λ in Lemma 4.1.25. Specifically, if we know that the quadrature weights are positive, then we can apply Lemma 4.2.12. It has been conjectured that for quasi-uniform sets of centers Λ, the quadrature weights are positive; however, this has yet to be proved. There is, however, some empirical evidence to support the conjecture, which we exhibit in Figure 4.2.10. The data comes from experiments on $SO(3)$ with kernels Φ_m that will be introduced in Chapter 5.

Figure 4.1: Plots of $\log(w_{\min})$ versus $\log(q_\Lambda)$, where $w_{\min}$ is the minimum quadrature weight for centers Λ, for samples Λ of $SO(3)$ of sizes 396, 896, 1005, 1872, and 3749.



As mentioned in Remark 4.1.26, we don't need a baseline oversampling exponent if we know the quadrature weights are positive. In that case we have a more direct way to get the result from Lemma 4.1.25. We need the following technical lemma for the proof.

Lemma 4.2.11. *Adopt the notation and assumptions from Lemma 4.1.14, suppose $u \in V_\Xi$, and let I_Λ be the interpolation operator relative to the centers Λ. There is a constant C such that, provided $h_\Lambda \leq C h_\Xi$,*

$$\frac{1}{2}\left\|u\right\|_{L_2} \leq \left\|I_\Lambda u\right\|_{L_2} \leq \frac{3}{2}\left\|u\right\|_{L_2}.$$

Proof. Two applications of the triangle inequality give

$$\|u\|_{L_2} - \|u - I_\Lambda u\|_{L_2} \leq \|I_\Lambda u\|_{L_2} \leq \|u\|_{L_2} + \|u - I_\Lambda u\|_{L_2}.$$

Thus, it suffices to show a C exists that ensures $\|u - I_\Lambda u\|_{L_2} \leq \frac{1}{2}\|u\|_{L_2}$. By our interpolation error estimate,

$$\|u - I_\Lambda u\|_{L_2} \leq C_1 h_\Lambda^m \|u\|_{H^m}.$$

By [14, Theorem 10], $\|u\|_{H^m} \leq C_2 h_\Xi^{-m}$ for $u \in V_\Xi$. Applying it to our situation gives

$$\|u - I_\Lambda u\|_{L_2} \leq C_1 C_2 \left(\frac{h_\Lambda}{h_\Xi}\right)^m \|u\|_{L_2}.$$

Hence, it suffices to take $C = (2C_1C_2)^{-1/m}$. □

Lemma 4.2.12. *Adopt the notation and assumptions from Lemma 4.1.14. Assume the quadrature weights $w = \{w_\zeta\}_{\zeta\in\Lambda}$ are positive, and let $w_{\min}$ be the minimum quadrature weight. There is a constant $C \geq h_\Lambda w_{\min}^{-1}$ such that, for h_Ξ sufficiently small,*

$$\left\|\left(B_\Xi^\Lambda\right)^{-1}\right\|_2 \leq C h_\Xi^{-d}.$$

Proof. As in the proof of Lemma 3.4.4, it suffices to show that $\lambda_{\min}\left(B_\Xi^\Lambda\right) \geq C q_\Xi^d$, and for that it suffices to show that $v^T B_\Xi^\Lambda v \geq C q_\Xi^d \|v\|_{\ell_2(\Xi)}^2$ for all vectors $v = \{v_\xi\}_{\xi\in\Xi}$. Also as in the proof of Lemma 3.4.4, we will need the L_2 stability of the Lagrange basis $\{\chi_\xi\}_{\xi\in\Xi}$, but now we will also need the L_2 stability of the Lagrange basis $\{\widetilde{\chi}_\zeta\}_{\zeta\in\Lambda}$; namely, that there are constants

D_1 and D_2 such that

$$D_1 q_\Lambda^{d/2} \|y\|_{\ell_2(\Lambda)} \le \left\| \sum_{\zeta\in\Lambda} y_\zeta \widetilde{\chi}_\zeta \right\|_{L_2} \le D_2 q_\Lambda^{d/2} \|y\|_{\ell_2(\Lambda)}$$

for all vectors $y = \{y_\zeta\}_{\zeta\in\Lambda}$. So, starting with a vector $v = \{v_\xi\}_{\xi\in\Xi}$, let $u = \sum_{\xi\in\Xi} v_\xi \chi_\xi$, so that $\nabla u = \sum_{\xi\in\Xi} v_\xi \nabla\chi_\xi$, and let $y = \{y_\zeta\}_{\zeta\in\Lambda} = u|_\Lambda$. We have

$$\begin{aligned}
v^T B_\Xi^\Lambda v &= \sum_{\xi\in\Xi}\sum_{\eta\in\Xi} v_\xi v_\eta \sum_{\zeta\in\Lambda} \left(a^\sharp(\zeta)\left(\nabla\chi_\xi(\zeta), \nabla\chi_\eta(\zeta)\right) + b(\zeta)\chi_\xi(\zeta)\chi_\eta(\zeta)\right) w_\zeta \\
&= \sum_{\zeta\in\Lambda} \left(a^\sharp(\zeta) \left(\sum_{\xi\in\Xi} v_\xi \nabla\chi_\xi(\zeta), \sum_{\eta\in\Xi} v_\eta \nabla\chi_\eta(\zeta) \right) \right. \\
&\qquad\qquad \left. + \;\; b(\zeta) \left(\sum_{\xi\in\Xi} v_\xi \chi_\xi(\zeta) \right) \left(\sum_{\eta\in\Xi} v_\eta \chi_\eta(\zeta) \right) \right) w_\zeta \\
&= \sum_{\zeta\in\Lambda} \left(a^\sharp(\zeta)\left(\nabla u(\zeta), \nabla u(\zeta)\right) + b(\zeta) u(\zeta)^2 \right) w_\zeta \\
&= Q^\Lambda \left(a^\sharp\left(\nabla u, \nabla u\right) + b u^2 \right).
\end{aligned}$$

Now, the positive definiteness of a ensures $a^\sharp(\nabla u, \nabla u) > 0$. Using this, the assumption that $w_\zeta \ge C h_\Lambda^d$ for each $\zeta \in \Lambda$, and the assumption that $b \ge b_0 > 0$ gives

$$\begin{aligned}
v^T B_\Xi^\Lambda v \ge Q^\Lambda\left(b u^2\right) &= \sum_{\zeta\in\Lambda} b(\zeta) u(\zeta)^2 w_\zeta \\
&\ge C b_0 h_\Lambda^d \sum_{\zeta\in\Lambda} u(\zeta)^2 = C b_0 h_\Lambda^d \|y\|^2_{\ell_2(\Lambda)}.
\end{aligned}$$

Let $\widetilde{u} = I_{\Lambda,L} u = \sum_{\zeta\in\Lambda} y_\zeta \widetilde{\chi}_\zeta$ be the interpolant to u on Λ. The L_2 stability of $\{\widetilde{\chi}_\zeta\}_{\zeta\in\Lambda}$ gives

$$v^T B_\Xi^\Lambda v \geq C b_0 h_\Lambda^d D_2^{-1} q_\Lambda^{-d} \|\widetilde{u}\|_{L_2}^2 = C b_0 D_2^{-1} \rho^d \|\widetilde{u}\|_{L_2}^2 .$$

By Lemma 4.2.11, $\|\widetilde{u}\|_{L_2}^2 \geq \frac{1}{4}\|u\|_{L_2}^2$, and this along with the L_2 stability of $\{\chi_\xi\}_{\xi\in\Xi}$ gives

$$v^T B_\Xi^\Lambda v \geq \frac{1}{4} C b_0 D_2^{-1} \rho^d \|u\|_{L_2}^2 \geq \frac{1}{4} C b_0 D_2^{-1} \rho^d C_1 q_\Xi^d \|v\|_{\ell_2(\Xi)}^2 .$$

□

4.3 Algorithms

To implement our Galerkin approximation numerically, we need to calculate the quadrature weights, the entries of the quadratized stiffness matrix, and the Galerkin approximation. We postpone the algorithm for the quadrature weights to Section 5.6, when we'll know more about the integrals of our kernel and the functions in our auxiliary space.

Suppose $\{b_1, \ldots, b_N\}$ is a basis for our set of centers $\Xi = \{\xi_1, \ldots, \xi_N\}$ for Galerkin approximation, let $\Lambda = \{\zeta_1, \ldots, \zeta_M\}$ be our set of centers for quadrature, and let $w = \{w_1, \ldots, w_M\}$ be the quadrature weights. We wish to form the quadratized stiffness matrix $S = [S_{ij}]_{i,j=1}^N$, with entries

$$\begin{aligned} S_{i,j} &= Q^\Lambda\left(a^\sharp\left(\nabla b_i, \nabla b_j\right) + b b_i, b_j\right) \\ &= \sum_{k=1}^M w_k \left(a^\sharp(\zeta_k)\left(\nabla b_i, \nabla b_j\right) + b(\zeta_k) b_i(\zeta_k) b_j(\zeta_k)\right) \\ &= \sum_{k=1}^M w_k \left(\sum_{m,n=1}^d a^{mn}(\zeta_k) \frac{\partial b_i}{\partial x^m}(\zeta_k) \frac{\partial b_j}{\partial x^n}(\zeta_k) + b(\zeta_k) b_i(\zeta_k) b_j(\zeta_k)\right). \end{aligned}$$

For a positive definite kernel Φ, one can simply take $b_j = \Phi(\cdot, \xi_j)$. The following algorithm computes the quadratized stiffness matrix, given the centers for quadrature, the quadrature weights, the coefficients in our differential equation, and a basis for V_Ξ.

Algorithm 1 Quadratized Stiffness Matrix for $\langle u, v\rangle_{a,b} = \lambda_f(v)$.

Input: $\Lambda = \{\zeta_1, \ldots, \zeta_M\}$, quadrature weights $\{w_1, \ldots, w_M\}$, the functions a^{ij} and b from our differential equation, and a basis $\{b_1, \ldots, b_N\}$ for V_Ξ.

Output: The quadratized stiffness matrix.

- Form the basis matrices

$$B = \begin{bmatrix} b_1(\zeta_1) & b_2(\zeta_2) & \cdot & b_N(\zeta_1) \\ b_1(\zeta_2) & b_2(\zeta_2) & \cdots & b_N(\zeta_2) \\ \vdots & \vdots & & \vdots \\ b_1(\zeta_M) & b_2(\zeta_M) & \cdots & b_N(\zeta_M) \end{bmatrix},$$

and

$$B_i = \begin{bmatrix} \frac{\partial}{\partial x^i} b_1(\zeta_1) & \frac{\partial}{\partial x^i} b_2(\zeta_1) & \cdots & \frac{\partial}{\partial x^i} b_N(\zeta_1) \\ \frac{\partial}{\partial x^i} b_1(\zeta_2) & \frac{\partial}{\partial x^i} b_2(\zeta_2) & \cdots & \frac{\partial}{\partial x^i} b_N(\zeta_2) \\ \vdots & \vdots & & \vdots \\ \frac{\partial}{\partial x^i} b_1(\zeta_M) & \frac{\partial}{\partial x^i} b_2(\zeta_M) & \cdots & \frac{\partial}{\partial x^i} b_N(\zeta_M) \end{bmatrix}$$

- Compute $D_{bw} = \mathrm{diag}(bw)$ and $D_{a^{ij}w} = \mathrm{diag}(a^{ij}w)$.
- Obtain the stiffness matrix $S = B^* D_{bw} B + \sum\limits_{i,j=1}^{d} B_i^* D_{a^{ij}w} B_j$.

For a conditionally positive definite kernel Φ, the functions $\Phi(\cdot, \xi_j)$ aren't even in V_Ξ. In this case one has to construct a basis using both the translates of the kernel and the functions in the auxiliary space Π. One can take the Lagrange basis, or take the following approach, due to Wendland - see [29, Section 10.3] for details. Choose a Π-unisolvent subset $\Xi_0 = \{x_1, \ldots, x_L\}$ of Ξ, set $\Xi_1 = \Xi \setminus \Xi_0 = \{y_1, \ldots, y_{N-L}\}$, and let $\{p_1, \ldots, p_L\}$ be the Lagrange basis for Π. Extend $\{p_1, \ldots, p_L\}$ to a basis $\{b_1, \ldots, b_{N-L}, p_1, \ldots, p_L\}$ by setting $b_j = \Phi(\cdot, y_j) - \sum_{k=1}^{L} p_k(y_j)\Phi(\cdot, x_k)$.

Algorithm 2 Basis for V_Ξ.

Input: kernel Φ, auxiliary space Π of dimension Q, and a Π-unisolvent set of centers $\Xi = \{\xi_1, \dots, \xi_N\}$.
Output: a basis $\{b_1, \dots, b_N\}$ for V_Ξ.
- Choose a (quasi-uniform) subset $\Xi_0 = \{x_1, \dots, x_Q\}$.
- Let $\Xi \setminus \Xi_0 = \Xi_1 = \{y_1, \dots, y_{N-Q}\}$.
- Form the Lagrange basis $\{p_1, \dots, p_Q\}$ for Ξ_0 and Π. (We can do this by forming the collocation matrix $K = \{\varphi_j(\xi_k)\}_{j,k=1}^{Q}$; then the coefficients of p_j are in the j^{th} column of K^{-1}.)
- For $1 \le j \le N - Q$ let

$$b_j = \Phi(\cdot, y_j) - \sum_{k=1}^{Q} p_k(y_j)\Phi(\cdot, x_k).$$

- For $N - Q + 1 \le j \le N$, let $b_j = p_{j-N+Q}$.
- Then $\{b_1, \dots, b_{N-Q}, p_1, \dots, p_Q\}$ is a basis for V_Ξ.

Once the basis and stiffness matrix has been formed, given a function f that is to be the right side of our differential equation, the Galerkin approximation is formed by pairing the basis elements with coefficients that are determined by a linear system. That is described in Algorithm 3.

Algorithm 3 Quadratized Galerkin Approximation to solution of $\mathcal{L}u = f$.

Input: $f|_\Lambda$, the stiffness matrix S, the quadrature weights w, and a basis matrix B as formed in Algorithm 4.3..
Output: the quadratized Galerkin approximation.
- Compute $D_{bw} = \operatorname{diag}(bw)$.
- Compute the RHS $R = B^* D_{bw}\, f|_\Lambda$.
- Solve $Sa = R$ for the coefficients a.
- The quadratized Galerkin approximation is $u_\Xi^\Lambda = \sum_{i=1}^{N} a_i b_i$.

Chapter 5

The Rotation Group

We now specialize to the particular manifold which we will be working on: the rotation group, $SO(3)$. It consists of all orthogonal 3×3 matrices with real entries and determinant 1:

$$SO(3) = \left\{ A \in \mathbb{R}^{3\times 3} : A^T A = I = AA^T, \det(A) = 1 \right\}.$$

In other words, it is the space of all proper rotations of $\mathbb{R}^3$. It is a compact Lie group of dimension 3. As such, it has a unique bi-invariant Haar measure, μ, and we assume that it is normalized; i.e. $\mu(SO(3)) = 1$.

5.1 Parametrizations

There are various parameterizations of $SO(3)$; we briefly present a few. The first is referred to as *Euler angles.* Euler showed that every $A \in SO(3)$ can be factored as

$$A = E(\varphi_1, \theta, \varphi_2) := R_z(\varphi_1) R_x(\theta) R_z(\varphi_2),$$

where $R_x(\alpha)$ and $R_z(\alpha)$ are rotations of angle α about the x- and z-axes, respectively:

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \quad R_z(\alpha) = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

This is the so-called ZXZ Euler angle decomposition. Other kinds exist, for example XYZ, YZY, etc. For rotations about the y-axis, we have

$$R_y(\alpha) = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}.$$

For a function $f : SO(3) \to \mathbb{R}$, the Haar integral can be rewritten as

$$\int_{SO(3)} f(x)d\mu(x) = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} f(\varphi_1, \theta, \varphi_2)\sin\theta d\varphi_1 d\theta d\varphi_2,$$

where we write $f(\varphi_1, \theta, \varphi_2)$ to mean $f(E(\varphi_1, \theta, \varphi_2))$. Euler angles, it should be noted, are not unique; for example, if $\theta = 0$ then $E(\varphi_1, \theta, \varphi_2) = R_z(\varphi_1 + \varphi_2)$ can be expressed in any number of ways.

Next, we have the *axis-angle* parametrization. Every element A of $SO(3)$ has 1 as an eigenvalue; i.e. an axis that it keeps fixed. A corresponding unit eigenvector v is called an *axis* for A. In the plane $v^\perp$, vectors rotated by A remain in $v^\perp$, and the angle by which they are rotated is called the *angle* of A, and denoted $\omega(A)$. A formula for the angle of a rotation A is

$$\omega(A) = \operatorname{acos}\left(\frac{\operatorname{tr}(A) - 1}{2}\right).$$

Axis-angle parameterizations are not unique; for example, taking the axis to be $-v$ and the angle to be $-\alpha$ also yields A. If $v = [v_1, v_2, v_3]^T$ and $\alpha \in (-\pi, \pi]$, then the matrix with axis span(v) and angle α is

$$A = \begin{bmatrix} v_1^2(1-c)+c & v_1v_2(1-c)+v_3s & v_1v_3(1-c)-v_2s \\ v_1v_2(1-c)-v_3s & v_2^2(1-c)+c & v_2v_3(1-c)+v_1s \\ v_1v_3(1-c)+v_2s & v_2v_3(1-c)-v_1s & v_3^2(1-c)+c \end{bmatrix},$$

where $c = \cos\alpha$ and $s = \sin\alpha$.

Our next parameterizations is via quaternions. A *quarternion* is a number of the form $w+x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$, where w, x, y, and z are real numbers and $\mathbf{i}$, $\mathbf{j}$, $\mathbf{k}$ are the "purely imaginary" quaternions, satisfying $\mathbf{i}^2 = \mathbf{j}^2 = \mathbf{k}^2 = -1$, $\mathbf{ij} = \mathbf{k}$, $\mathbf{jk} = \mathbf{i}$, $\mathbf{ki} = \mathbf{j}$, $\mathbf{ji} = -\mathbf{k}$, $\mathbf{kj} = -\mathbf{i}$, and $\mathbf{ik} = -\mathbf{j}$. There is a 2-to-1 homomorphism from the group of unit quaternions to $SO(3)$, a unit quaternion being a quaternion $w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ for which $w^2 + x^2 + y^2 + z^2 = 1$. The unit quaternion $w + x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$ maps to the matrix

$$A = \begin{bmatrix} 1-2y^2-2z^2 & 2xy-2wz & 2xy-2wy \\ 2xy+2wz & 1-2x^2-2z^2 & 2yz-2wx \\ 2xz-2wy & 2yz+2wx & 1-2x^2-2y^2 \end{bmatrix}.$$

This may seem to suggest that $SO(3)$ is 4-dimensional, but, the set of unit quaternions is in fact 3-dimensional, as it naturally identifies with the unit sphere in $\mathbb{R}^4$.

The last parameterizations we will see is via the matrix exponential. The Lie algebra of $SO(3)$ is $so(3)$, the space of real skew-symmetric matrices.

If we set

$$X_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, X_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}, X_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

then every $X \in so(3)$ can be expressed as $\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3$ with $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$. The *matrix exponential*, $\exp : so(3) \to SO(3)$ is an isometry, where in general for a 3×3 matrix X

$$\exp(X) = \sum_{n=0}^{\infty} \frac{1}{n!} X^n.$$

Direct calculations show that $\exp(\alpha X_1) = R_x(\alpha)$, $\exp(\alpha X_2) = R_y(\alpha)$, and $\exp(\alpha X_3) = R_z(\alpha)$; however, since X_1, X_2, and X_3 don't commute, calculating $\exp(\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3)$ is complicated in general. If, on the other hand, if $\mathbf{u} = [\alpha_1, \alpha_2, \alpha_3]^T$ is a unit vector in $\mathbb{R}^3$ and we let $X = \alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3$, then for $\theta \in \mathbb{R}$, $\exp(\theta X)$ has axis $\mathbf{u}$ and angle θ, and the Rodrigues addition formula gives

$$\exp(\theta X) = I + \sin\theta X + (1 - \cos\theta) X^2,$$

which is where the parameterizations above for axis-angle comes from.

5.2 Harmonic Analysis

Much of the material in this section can be found in [26] for the specific case of $SO(3)$ and in [10] for the general case. Being a compact group, the Peter-Weyl Theorem (see [26, Theorem 2.24] or [10, Theorem 5.12], for example) guarantees that the matrix elements of the irreducible unitary representations of $SO(3)$ form a complete orthogonal set for $L_2(SO(3))$. These are the *Wigner-D* functions: for each $\ell \in \mathbb{N}$ there is a unique (up to unitary equivalence) unitary representation D^ℓ which has dimension $2\ell+1$. For $j, k \in \{-\ell, \ldots, \ell\}$ the Wigner-D functions are the matrix elements $D^\ell_{j,k}$ of D^ℓ. A different (unitarily equivalent) choice for D^ℓ will produce different Wigner-D functions, but their span will be the same.

There are many different formulations for the Wigner-D functions $D^\ell_{j,k}$. One of the more common can be found in [26] and [17], given in Euler angles. They are:

$$D^\ell_{j,k}(\varphi_1, \theta, \varphi_2) = e^{-j\varphi_1} P^\ell_{j,k}(\cos\theta) e^{-ik\varphi_2},$$

where $P^\ell_{j,k}$ is given by

$$P^\ell_{j,k}(t) = C(1-t)^{-\frac{k-j}{2}}(1+t)^{-\frac{j+k}{2}} \frac{d^{\ell-k}}{dt^{\ell-k}} \left[(1-t)^{\ell-j}(1+t)^{\ell+j}\right]$$

and $C = \frac{(-1)^{\ell-j} i^{k-j}}{2^\ell(\ell-j)!} \sqrt{\frac{(1-j)!(\ell+k)!}{(\ell+j)!(\ell-k)!}}$. In this formulation, the Wigner-D functions are complex-valued.

There are also real-valued Wigner-D functions found in [20], given by

$$\begin{aligned} D^\ell_{j,k}(\varphi_1, \theta, \varphi_2) = \operatorname{sgn}(k)\Psi_j(\varphi_1)\Psi_k(\varphi_2) &\frac{d^\ell_{|k|,|j|}(\theta) + (-1)^j d^\ell_{|j|,-|k|}(\theta)}{2} \\ - \operatorname{sgn}(j)\Psi_{-j}(\varphi_1)\Psi_{-k}(\varphi_2) &\frac{d^\ell_{|k|,|j|}(\theta) - (-1)^j d^\ell_{|j|,-|k|}(\theta)}{2}, \end{aligned}$$

where

$$\Psi(x) = \begin{cases} \sqrt{2}\cos(jx) & \text{if } m > 0 \\ 1 & \text{if } m = 0 \\ \sqrt{2}\sin(|j|x) & \text{if } m < 0 \end{cases}$$

and the little Wigner-D functions are

$$d^{\ell}_{j,k}(t) = \sqrt{\frac{(\ell+k)!(\ell-k)!}{(\ell+j)!(\ell-j)!}} \left(\sin\frac{t}{2}\right)^{k-j} \left(\cos\frac{t}{2}\right)^{j+k} P^{(\ell-k)}_{(k-j,j+k)}(\cos t),$$

with $P^{(\cdot)}_{(\cdot,\cdot)}$ the Jacobi polynomials.

For each $\ell \in \mathbb{N}$ and $j,k \in \{-\ell,\dots,\ell\}$, $D^{\ell}_{j,k}$ is an eigenfunction of $-\Delta$ corresponding to the eigenvalue $\lambda_\ell = \ell(\ell+1)$. We refer to ℓ as the *degree* of $D^{\ell}_{j,k}$ and j and k as the *orders*. We denote by Π_L the set of Wigner-D functions of degree L or less. These are our auxiliary spaces for interpolation with conditionally positive definite kernels on $SO(3)$.

5.3 Orthonormal Basis Eigenfunctions

The Wigner-D functions are orthogonal in L_2, but not orthonormal. In order to be able to express our kernels in the form (2.2.1), we must take

$$\varphi^{\ell}_{j,k} = \frac{D^{\ell}_{j,k}}{\left\|D^{\ell}_{j,k}\right\|_{L_2}} = \sqrt{2\ell+1}D^{\ell}_{j,k}.$$

In fact, it is part of the Peter-Weyl Theorem that these form an orthonormal basis for L_2. For a function $f \in L_2(SO(3))$, $\ell \in \mathbb{N}$ and $j,k \in \{-\ell,\dots,\ell\}$, if we define

$$\widehat{f}^{\ell}_{j,k} = \int_{SO(3)} f\varphi^{\ell}_{j,k}d\mu = \sqrt{2\ell+1}\int_{SO(3)} fD^{\ell}_{j,k}d\mu,$$

then

$$f = \sum_{\ell=0}^{\infty}\sum_{j,k=-\ell}^{\ell} \widehat{f}^{\ell}_{j,k}\varphi^{\ell}_{j,k} = \sum_{\ell=0}^{\infty}\sqrt{2\ell+1}\sum_{j,k=-\ell}^{\ell}\widehat{f}^{\ell}_{j,k}D^{\ell}_{j,k}.$$

Of particular interest to us are functions that depend only on the rotation angle, called *class functions*. These are spanned by the *characters* $\mathfrak{c}_\ell = \operatorname{tr}\left(D^\ell\right)$. Thus, for a class function $f : SO(3) \to \mathbb{R}$ there is a uniquely determined $\widetilde{f} : [0, \pi] \to \mathbb{R}$ for which $f(x) = \widetilde{f}(\omega(x))$ for all $x \in SO(3)$. For such functions the Haar integral simplifies to

$$\int_{SO(3)} f(x)d\mu(x) = \frac{2}{\pi}\int_0^\pi \widetilde{f}(t)\sin^2\left(\frac{t}{2}\right) dt.$$

We also have the following addition formula for Wigner-D functions:

$$\begin{aligned}\sum_{j,k=-\ell}^{\ell} D^\ell_{j,k}(x)D^\ell_{j,k}(y) &= \operatorname{tr}\left(D^\ell(x)D^\ell(y^{-1})\right) = \operatorname{tr}\left(D^\ell\left(y^{-1}x\right)\right) \\ &= \mathfrak{c}_\ell\left(y^{-1}x\right) = \mathcal{U}_{2\ell}\left(\cos\left(\frac{\omega\left(y^{-1}x\right)}{2}\right)\right),\end{aligned}$$

where $\mathcal{U}_n$ is the Chebyshev polynomial of the second kind of degree n. The representation corresponding to our complete orthonormal set is $\varphi^\ell = \sqrt{2\ell+1}D^\ell$, and the addition formula for them is

$$\sum_{j=-\ell}^{\ell} \varphi^\ell_{j,k}(x)\varphi^\ell_{j,k}(y) = \sqrt{2\ell+1}\,\mathcal{U}_{2\ell}\left(\cos\left(\frac{\omega\left(y^{-1}x\right)}{2}\right)\right).$$

We finish this section with lemmas that give the H^m norm of our normalized eigenfunctions and their covariant derivatives.

Lemma 5.3.1. *For* $\tau \in \mathbb{N}$, $\left\|\varphi^\ell_{j,k}\right\|_{H^\tau} = (1+\lambda_\ell)^{\tau/2}$.

Proof.

$$\begin{aligned}\left\|\varphi^\ell_{j,k}\right\|^2_{H^\tau} &= \left\langle\varphi^\ell_{j,k}, \varphi^\ell_{j,k}\right\rangle_{H^\tau} = \left\langle(I-\Delta)^\tau\varphi^\ell_{j,k}, \varphi^\ell_{j,k}\right\rangle_{L_2} \\ &= (1+\lambda_\ell)^\tau\left\langle\varphi^\ell_{j,k}, \varphi^\ell_{j,k}\right\rangle_{L_2} = (1+\lambda_\ell)^\tau.\end{aligned}$$

□

Lemma 5.3.2. *For $\tau \in \mathbb{N}$, $\left\|\nabla \varphi_{j,k}^{\ell}\right\|_{H^\tau} \leq C\ell^{\tau+1}$.*

Proof. By Lemma 2.5.3, there are numbers a_n, $n = 0, \ldots 2\tau$ such that $(I - \Delta)^\tau = \sum_{n=0}^{2\tau} a_n (\nabla^n)^* \nabla^n$. Thus,

$$\begin{aligned}
\left\|\nabla \varphi_{j,k}^{\ell}\right\|_{H^\tau}^2 &= \left\langle \nabla \varphi_{j,k}^{\ell}, \varphi_{j,k}^{\ell} \right\rangle_{H^\tau} = \left\langle (I-\Delta)^\tau \nabla \varphi_{j,k}^{\ell}, \nabla \varphi_{j,k}^{\ell} \right\rangle_{L_2} \\
&= \sum_{n=0}^{2\tau} a_n \left\langle \nabla^{n+1} \varphi_{j,k}^{\ell}, \nabla^{n+1} \varphi_{j,k}^{\ell} \right\rangle_{L_2} \\
&= \sum_{n=0}^{2\tau} a_n \left\langle \left(\nabla^{n+1}\right)^* \nabla^{n+1} \varphi_{j,k}^{\ell}, \varphi_{j,k}^{\ell} \right\rangle_{L_2}.
\end{aligned}$$

By Lemma 2.5.2 there are numbers b_i, $i = 0, \ldots 2\tau + 1$ such that

$$\sum_{n=0}^{2\tau} a_n \left(\nabla^{n+1}\right)^* \nabla^{n+1} = \sum_{i=0}^{2\tau+1} b_i \Delta^i.$$

Thus,

$$\begin{aligned}
\left\|\nabla \varphi_{j,k}^{\ell}\right\|_{H^\tau}^2 &= \sum_{i=0}^{2\tau+1} b_i \left\langle \Delta^i \varphi_{j,k}^{\ell}, \varphi_{j,k}^{\ell} \right\rangle_{L_2} \\
&= \sum_{i=0}^{2\tau+1} b_i (-1)^i \lambda_\ell^i \left\langle \varphi_{j,k}^{\ell}, \varphi_{j,k}^{\ell} \right\rangle_{L_2} \\
&\leq C\ell^{2\tau+2}
\end{aligned}$$

where we have taken $C = \max_{i \in \{1,\ldots,2\tau+1\}} |b_i|$. Taking square roots yields the result. □

5.4 Kernels and Spaces

We recall here certain definitions from Chapters 2 and 4, restated using notation better suited for $SO(3)$. We now *do* assume that the eigenvalues λ_ℓ

are distinct; our Hilbert-Schmidt series will have a form similar to (2.2.1), except now we group the basis eigenfunctions for the same eigenvalue together. Thus, our kernels $\Phi : SO(3) \times SO(3) \longrightarrow \mathbb{R}$ have the form

$$\Phi(x, y) = \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell) \sum_{j,k=-\ell}^{\ell} \varphi_{j,k}^{\ell}(x)\varphi_{j,k}^{\ell}(y). \tag{5.4.1}$$

We assume the coefficients satisfy (2.8.1) for some $\tau \in \mathbb{N}$ with $\tau > 3/2$, either for all $\ell \in \mathbb{N}$ if Φ is positive definite, or for all $\ell > L$ if Φ is conditionally positive definite with respect to Π_L. This ensures that the native space norm is equivalent to the Sobolev norm in the positive definite case.

The native space for Φ is

$$\mathcal{N}_{\Phi} = \left\{ f \in L_2 : \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)^{-1} \sum_{j,k=-\ell}^{\ell} \left| \widehat{f}_{j,k}^{\ell} \right|^2 < \infty \right\},$$

with inner product

$$\langle f, g \rangle_{\mathcal{N}_{\Phi}} = \sum_{\ell=0}^{\infty} \widehat{\phi}(\ell)^{-1} \sum_{j,k=-\ell}^{\ell} \widehat{f}_{j,k}^{\ell} \overline{\widehat{g}_{j,k}^{\ell}}.$$

The Sobolev space H^{τ} is

$$H^{\tau} = \left\{ f \in L_2 : \sum_{\ell=0}^{\infty} (1 + \lambda_{\ell})^{\tau} \sum_{j,k=-\ell}^{\ell} \left| \widehat{f}_{j,k}^{\ell} \right|^2 < \infty \right\},$$

with inner product

$$\langle f, g \rangle_{H^{\tau}} = \sum_{\ell=0}^{\infty} (1 + \lambda_{\ell})^{\tau} \sum_{j,k=-\ell}^{\ell} \widehat{f}_{j,k}^{\ell} \overline{\widehat{g}_{j,k}^{\ell}}.$$

For a subset $\Omega \subset SO(3)$, we also have the Sobolev space

$$W_2^m(\Omega) = \left\{ f \in L_2(\Omega) : \sum_{k=0}^{m} \int_\Omega \left\langle \nabla^k f, \nabla^k f \right\rangle_p d\mu(p) < \infty \right\},$$

with inner product

$$\langle f, g \rangle_{W_2^m(\Omega)} = \sum_{k=0}^{m} \int_\Omega \left\langle \nabla^k f, \nabla^k g \right\rangle_p d\mu(p).$$

Since $SO(3)$ is compact and without boundary, when Ω is all of $SO(3)$ and $\tau = m$ is an integer, $H^m = W_2^m$ with the norms being equivalent by Proposition 2.7.3. If in addition (2.8.1) holds for all $\ell \in \mathbb{N}$; i.e. Φ is positive definite, then $\mathcal{N}_\Phi = H^m = W_2^m$, with all three norms being equivalent.

5.5 Energy Estimates: Two Examples

We now exhibit two examples of kernels on $SO(3)$ that provide energy estimates. They both have advantages and drawbacks. The first has the advantage that a closed-form formula for it is known, but has the drawback that it only provides an algebraic energy estimate. The second has the advantage that it provides an exponential energy estimate, but has the drawback that no closed-form formula for it is known.

Example 5.5.1. An example of a family of kernels of the form (5.4.1) are the *rotational surface splines*: for $m \geq 3$,

$$\Phi_m(x, y) = \left(\sin\left(\frac{\omega\left(y^{-1}x\right)}{2} \right) \right)^{2m-3}.$$

The coefficients of Φ_m, found in [17], are

$$\widehat{\phi}_m(\ell) = \frac{(2m-2)!}{\pi(-4)^{m-1}}(2\ell+1)\prod_{j=0}^{m-1}\frac{1}{\ell(\ell+1)-\left(j^2-\frac{1}{4}\right)}.$$

and Φ_m is conditionally positive definite with respect to Π_{m-2}. The coefficients $\widehat{\phi}_m(\ell)$ satisfy (2.8.1) with $\tau = m - \frac{1}{2}$. The corresponding differential operator, for which Φ_m is the fundamental solution, is

$$\mathcal{L} = \frac{(2m-2)!}{\pi(-4)^{m-1}}\prod_{j=0}^{m-1}\left(\Delta - \left(j^2 - \frac{1}{4}\right)\right).$$

This operator does *not* annihilate Π_{m-2}, and so Φ_m provides only an algebraic energy estimate.

Let Ξ and Λ be Π_{m-2}-unisolvent sets of centers with $\#\Xi < \#\Lambda$ and $h_\Lambda < h_\Xi$, and consider using Φ_m to produce the quadratized Galerkin approximation u_Ξ^Λ to the weak solution of $\mathcal{L}u = f$, with $f \in H^m$, using Ξ as the set of centers for Galerkin approximation and Λ the set of centers for quadrature. Then Corollary 4.2.7 guarantees that

$$\left\|u - u_\Xi^\Lambda\right\|_{L_2} \leq C\left(h_\Xi^{m-1} + \left(\frac{h_\Lambda}{h_\Xi}\right)^m h_\Xi^{-11/2}h_\Lambda^{-5/2}\right)\|f\|_{H^s},$$

where s is determined as in Theorem 4.2.1, and Remark 4.2.8 shows we should take our oversampling exponent to be $p = 2 + \frac{19}{2m-5}$.

Example 5.5.2. Another example of a family of kernels of the form (5.4.1) are the "ideal" kernels

$$\kappa_m(x,y) = \sum_{\ell=0}^{\infty}(1+\lambda_\ell)^{-m}\sum_{j,k=-\ell}^{\ell}\varphi_{j,k}^\ell(x)\varphi_{j,k}^\ell(y).$$

Note that we have not written $\overline{\varphi_{j,k}^\ell(y)}$ - we assume from here on out that the Wigner-D functions, and hence our orthonormal eigenfunctions, are real-valued. Here (2.8.1) holds with $C_1 = C_2 = 1$ and $\tau = m$, since the coefficients of κ_m are exactly $\widehat{\kappa}_m(\ell) = (1+\lambda_\ell)^{-m}$. The corresponding operator $\mathcal{L}$ for which κ_m is the fundamental solution is $\mathcal{L} = (I-\Delta)^m$. Since $\widehat{\kappa}_m(\ell)$ is positive for all ℓ, κ_m is positive definite, and therefore provides an exponential energy estimate.

Let Ξ, Λ, and $f \in H^m$ be as in Example 5.5.1. If u_Ξ^Λ is the quadratized Galerkin approximation to the solution of $\mathcal{L}u = f$, then Corollary 4.2.7 guarantees that

$$\left\|u - u_\Xi^\Lambda\right\|_{L_2} \leq C\left(h_\Xi^{m-1} + \left(\frac{h_\Lambda}{h_\Xi}\right)^m h_\Xi^{-11/2} h_\Lambda^{-1} \left|\log\left(h_\Lambda\right)\right|^3\right) \|f\|_{H^m},$$

and Remark 4.2.8 shows that we should take our oversampling exponent to be $p = \frac{3}{\log(m)} W\left(\frac{1}{3}\log(m) h_\Xi^{m+1} \left|\log\left(h_\Xi\right)\right|^{-1}\right)$. Again, taking the oversampling exponent that would arise in the algebraic case will suffice, so, we can actually take $p = 2 + \frac{19}{2m-5}$.

Although the kernel from Example 5.5.2 doesn't have a closed form, we can still use it. In the next chapter, we show how to truncate it to get a Galerkin approximation. What we must determine, then, is *when* to truncate.

5.6 Algorithm: Quadrature Weights

To implement our Galerkin approximation scheme, we need one final algorithm, one for computing the quadrature weights. We restrict our attention to the kernel Φ_m from Section 5.5. As per Section 3.1, we need to know the integrals $\int_{SO(3)} \Phi_m(\cdot, y) d\mu$ and $\int_{SO(3)} \varphi_{j,k}^\ell d\mu$. The fact that the former integral is independent of y follows from the fact that Φ_m has a Hilbert-

Schmidt decomposition and the orthonormality of the eigenfunctions of the Laplace-Beltrami operator. Hence, we can take $y = \text{Id}$ to obtain

$$J = \int_{SO(3)} \Phi(\cdot, \text{Id}) d\mu = \frac{2}{\pi} \int_0^\pi \left(\sin\left(\frac{\omega}{2}\right)\right)^{2m-1} d\omega = \frac{2\Gamma(m)}{\sqrt{\pi}\Gamma(m+\frac{1}{2})}.$$

When $\ell = j = k = 0$, $\varphi_{0,0}^0 = 1$, and so the normalization of the Haar measure gives $\int_{SO(3)} \varphi_{0,0}^0 d\mu = 1$. For any other eigenfunction $\varphi_{j,k}^\ell$, orthonormality gives

$$J_{j,k}^\ell = \int_{SO(3)} \varphi_{j,k}^\ell d\mu = \left\langle \varphi_{j,k}^\ell, \varphi_{0,0}^0 \right\rangle_{L_2} = 0.$$

This leads us to the following.

Algorithm 4 Quadrature Weights

Input: $\Lambda = \{\zeta_1, \ldots, \zeta_M\}$, kernel parameter m.
Output: Quadrature weights $w = \{w_1, \ldots, w_M\}$.
• Form the auxiliary matrix

$$P = \left[\varphi_{j,k}^\ell\big|_\Lambda\right]_{\ell\in\{0,\ldots,m-2\}, j,k\in\{-\ell,\ell\}}.$$

• Form the collocation matrix K_Ξ.
• solve the system

$$\begin{bmatrix} K_\Xi & P \\ P^T & 0 \end{bmatrix} \begin{bmatrix} w \\ v \end{bmatrix} = \begin{bmatrix} J\mathbf{1} \\ u \end{bmatrix},$$

where $J = \frac{2\Gamma(m)}{\sqrt{\pi}\Gamma(m+\frac{1}{2})}$, $\mathbf{1}$ is the $M \times 1$ vector of 1's, and u is the $Q \times 1$ vector whose first entry is 1 and all other entries are 0 ($Q = \binom{2m-1}{3}$ is the dimension of the auxiliary space Π_{m-2}).
• The quadrature weights are w.

Chapter 6

A Truncated $SO(3)$ Kernel

The rotational surface splines have the advantage of having a closed form. The same can not be said for some other kernels on $SO(3)$, such as the one in Example 5.5.2. In practice, to use those kinds of kernels, some sort of truncation of the uniformly convergent series in (5.4.1) is desirable. We explore in this chapter the error that arises when such truncation occurs, in a specific case.

Our strategy is as follows. We employ a particular kernel κ_m, $m > 5/2$ with an expansion as in (5.4.1), and which provides an exponential energy estimate. We then truncate by taking only finitely many terms in (5.4.1) to get the truncated kernel $\widetilde{\kappa}_m$, the number of terms being N, the *truncation parameter*. We then form the truncated Lagrange basis $\{\widetilde{\chi}_\xi\}_{\xi\in\Xi}$, consisting of the Lagrange functions for the truncated kernel. The truncated stiffness matrix in the truncated Lagrange basis, $\widetilde{B}_\Xi$ is then used to obtain the truncated Galerkin approximation, $\widetilde{u}_\Xi$.

Upon establishing an L_2 truncation error estimate for the truncated Galerkin approximation, we then explore the error that arises by quadratizing the truncated Galerkin approximation. This means replacing the entries of the truncated stiffness matrix with quadrature estimates to form the quadratized truncated stiffness matrix $\widetilde{B}_{\Xi}^{\Lambda}$, which is then used to obtain the quadratized truncated Galerkin approximation, $\widetilde{u}_{\Xi}^{\Lambda}$. The full error estimate is then attained via repeated applications of the triangle inequality:

$$\left\|u - \widetilde{u}_{\Xi}^{\Lambda}\right\|_{L_2} \leq \|u - u_{\Xi}\|_{L_2} + \|u_{\Xi} - \widetilde{u}_{\Xi}\|_{L_2} + \left\|\widetilde{u}_{\Xi} - \widetilde{u}_{\Xi}^{\Lambda}\right\|_{L_2}.$$

We restrict our attention to the particular family of kernels from Example 5.5.2; namely,

$$\kappa_m(x, y) = \sum_{\ell=0}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j,k=-\ell}^{-\ell} \varphi_{j,k}^{\ell}(x)\varphi_{j,k}^{\ell}(y),$$

with Fourier-Chebyshev coefficients

$$\widehat{\kappa}_m(\ell) = (1 + \lambda_\ell)^{-m} = (1 + \ell(\ell + 1))^{-m}.$$

Note that we have not written the conjugate of $\varphi_{j,k}^{\ell}(y)$ in the expansion of our kernel. This is deliberate - to simplify matters, we assume throughout that the Wigner-D functions, and hence our orthonormal basis functions, are real-valued. Also note that the bound on the coefficients of κ_m,

$$|\widehat{\kappa}_m(\ell)| \leq C\ell^{-2m}, \tag{6.0.1}$$

holds for all $\ell \geq 1$.

Since $\widehat{\kappa}(\ell) > 0$ for all $\ell \in \mathbb{N}$, κ_m is positive definite. As such it provides an exponential energy estimate, and all our previous results for such kernels apply. Let ν be the constant from the energy estimate. In what follows whenever we say "there exists a positive constant C", it is understood that

C depends on the kernel Φ_m, the mesh ratio ρ of our centers Ξ, the tensor a and function b in our differential equation $\mathcal{L}$, and ν.

6.1 The Truncated Kernel

The *truncated kernel*, $\widetilde{\kappa}_m$, with *truncation parameter* N, is

$$\widetilde{\kappa}_m(x,y) = \sum_{\ell=0}^{N} \widehat{\kappa}_m(\ell) \sum_{j,k=-\ell}^{\ell} \varphi_{j,k}^{\ell}(x)\varphi_{j,k}^{\ell}(y).$$

Our first order of business is to establish pointwise and L_2 error estimates for the truncated kernel, and L_2 error estimates for its covariant derivative. We point out that both κ_m and $\widetilde{\kappa}_m$, and the covariant derivatives of both, enjoy an L_2-invariance in the sense that $\|\kappa_m(\cdot,x)\|_{L_2} = \|\kappa_m(\cdot,y)\|_{L_2}$ for all $x, y \in SO(3)$.

Lemma 6.1.1. *For $m > 1$ and $N > 0$ the uniform bound*

$$|\kappa_m(x,y) - \widetilde{\kappa}_m(x,y)| \leq CN^{2-2m}.$$

holds for all $x, y \in SO(3)$.

Proof. Since

$$\max_{\theta\in[0,\pi]} \left| \mathcal{U}_{2\ell}\left(\cos\left(\frac{\theta}{2}\right)\right)\right| = \mathcal{U}_{2\ell}(1) = 2\ell + 1,$$

we have by (6.0.1) that

$$
\begin{aligned}
|\kappa_m(x,y) - \widetilde{\kappa}_m(x,y)| &= \left| \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \mathcal{U}_{2\ell} \left(\cos\left(\frac{\omega\left(y^{-1}x\right)}{2} \right) \right) \right| \\
&\leq C \sum_{\ell=N+1}^{\infty} \ell^{-2m}(2\ell+1) \\
&\leq C' \sum_{\ell=N+1}^{\infty} \ell^{1-2m}.
\end{aligned}
$$

The result now follows from an integral comparison. □

Lemma 6.1.2. *For $m > 3/2$, $\tau \in \mathbb{N}$ with $0 \leq \tau < 2m-2$, $N > 0$, and any $\zeta \in SO(3)$,*

$$\|\kappa_m(\cdot,\zeta) - \widetilde{\kappa}_m(\cdot,\zeta)\|_{H^\tau} \leq CN^{2+\tau-2m}.$$

Proof. We need only prove the case $\zeta = \mathrm{Id}$. We have

$$
\begin{aligned}
\kappa_m(x,\mathrm{Id}) - \widetilde{\kappa}_m(x,\mathrm{Id}) &= \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j,k=-\ell}^{\ell} \varphi^{\ell}_{j,k}(x)\varphi^{\ell}_{j,k}(\mathrm{Id}) \\
&= \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j=-\ell}^{\ell} \varphi^{\ell}_{j,j}(x).
\end{aligned}
$$

Using (6.0.1) and Lemma 5.3.1, then, we have

$$
\begin{aligned}
\|\kappa_m(\cdot,\mathrm{Id}) - \widetilde{\kappa}_m(\cdot,\mathrm{Id})\|_{H^\tau} &\leq \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j=-\ell}^{\ell} \left\|\varphi^{\ell}_{j,j}\right\|_{H^\tau} \\
&\leq C \sum_{\ell=N+1}^{\infty} \ell^{\tau-2m}(2\ell+1) \\
&\leq C \sum_{\ell=N+1}^{\infty} \ell^{1+\tau-2m}.
\end{aligned}
$$

The result now follows from an integral comparison. □

Lemma 6.1.3. *For $m > 3/2$, $\tau \in \mathbb{N}$ with $0 \le \tau < 2m-3$, $N > 0$, and any $\zeta \in SO(3)$,*

$$\left\| \nabla \left(\kappa_m(\cdot, \zeta) - \widetilde{\kappa}_m(\cdot, \zeta) \right) \right\|_{H^\tau} \le C N^{3+\tau-2m}.$$

Proof. Again, we need only prove the case $\zeta = \mathrm{Id}$. We have

$$\begin{aligned} \nabla \left(\kappa_m(x, \mathrm{Id}) - \widetilde{\kappa}_m(x, \mathrm{Id}) \right) &= \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j,k=-\ell}^{\ell} \left(\nabla \varphi_{j,k}^{\ell}(x) \right) \varphi_{j,k}^{\ell}(\mathrm{Id}) \\ &= \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j=-\ell}^{\ell} \nabla \varphi_{j,j}^{\ell}(x). \end{aligned}$$

and so by (6.0.1) and Lemma 5.3.2,

$$\begin{aligned} \left\| \nabla \left(\kappa_m(\cdot, \mathrm{Id}) - \widetilde{\kappa}_m(\cdot, \mathrm{Id}) \right) \right\|_{H^\tau} &\le \sum_{\ell=N+1}^{\infty} \widehat{\kappa}_m(\ell) \sum_{j=-\ell}^{\ell} \left\| \nabla \varphi_{j,j}^{\ell} \right\|_{H^\tau} \\ &\le C \sum_{\ell=N+1}^{\infty} \ell^{1+\tau-2m}(2\ell+1) \\ &\le C \sum_{\ell=N+1}^{\infty} \ell^{2+\tau-2m}. \end{aligned}$$

The result now follows from an integral comparison. □

6.2 The Truncated Lagrange Basis

To get the *truncated Lagrange basis*, $\{\widetilde{\chi}_\xi\}_{\xi \in \Xi}$, we take the *truncated collocation matrix*, $\widetilde{K}_\Xi = \{\widetilde{\kappa}_m(\xi, \eta)\}_{\xi \eta \in \Xi}$, and the coefficients $\widetilde{\alpha}_\xi = \{\widetilde{\alpha}_{\xi,\eta}\}_{\eta \in \Xi}$ of $\widetilde{\chi}_\xi = \sum_{\eta \in \Xi} \widetilde{\alpha}_{\xi,\eta} \widetilde{\kappa}_m(\cdot, \eta)$ are obtained by solving the linear system $\widetilde{K} \widetilde{\alpha}_\xi = \delta_\xi$. We reiterate that the choice of basis does not affect interpolants, Galerkin approximations, or quadratized Galerkin approximation. It is used purely for the stability of the approximation scheme.

The first result in this section concerns the 1-norm of the inverse of the collocation matrix. In this result and in many that follow, we are computing the 1-norm of a self-adjoint matrix. We note that for a self-adjoint matrix A, $\|A\|_2 \leq \|A\|_1$, and so each such result has a corresponding corollary for the 2-norm.

Lemma 6.2.1. *Let K_Ξ be the collocation matrix for interpolation with the kernel κ_m on the set of centers Ξ, with $m > 3/2$ and mesh ratio ρ. There exists a constant C such that, for h_Ξ sufficiently small,*

$$\left\|K_\Xi^{-1}\right\|_1 \leq C h_\Xi^{3-2m}.$$

Proof. The entries in column ξ of K_Ξ^{-1} are precisely the coefficients of the Lagrange function $\chi_\xi = \sum_{\eta\in\Xi} \alpha_{\xi\eta}\kappa(\cdot,\eta)$. Hence,

$$\left\|K_\Xi^{-1}\right\|_1 = \max_{\xi\in\Xi} \sum_{\eta\in\Xi} \left|\alpha_{\xi\eta}\right|.$$

We follow [14] by noting that the proof of [11, Equation 5.6] gives, "mutatis mutandis",

$$|\alpha_{\xi\zeta}| \leq C q_\Xi^{3-2m} \exp\left(-\nu \frac{\operatorname{dist}(\xi,\zeta)}{h_\Xi}\right)$$

with positive constants C and ν that depend only on k_m. These are the same as the constants from our energy estimate. We can proceed by dividing $SO(3)$ into annuli A_n with center ξ, outer radius nh_Ξ, and inner radius $(n-1)h_\Xi$, $n = 1, 2, \ldots, n_{\max}$. We then have

$$\sum_{\zeta\in\Xi} |\alpha_{\xi\zeta}| = \sum_{n=1}^{n_{\max}} \sum_{\zeta\in A_n} |\alpha_{\xi\zeta}| \leq C q_\Xi^{3-2m} \sum_{n=1}^{n_{\max}} \sum_{\zeta\in A_n} \exp\left(-\nu \frac{\operatorname{dist}(\xi,\zeta)}{h_\xi}\right).$$

Each A_n has volume $\sim n^2 h_\Xi^3$, and so

$$\#\left(\Xi \cap A_n\right) \sim \frac{n^2 h_\Xi^3}{q_\Xi^3} = \rho^3 n^2.$$

Also, the minimum distance between $\zeta \in A_n$ and ξ is bounded below by $(n-1)h_\Xi$, which gives

$$\begin{aligned}\sum_{\zeta\in\Xi} |\alpha_{\xi\zeta}| &\leq C q_\Xi^{3-2m} \sum_{n=1}^{n_{\max}} \#\left(\Xi \cap A_n\right) \exp\left(-\nu \frac{(n-1)h_\Xi}{h_\Xi}\right) \\ &\leq C q_\Xi^{3-2m} \sum_{n=1}^{n_{\max}} n^2 e^{-\nu(n-1)} \\ &\leq C q_\Xi^{3-2m} \sum_{n=0}^{\infty} (n+1)^2 e^{-\nu n}.\end{aligned}$$

Summing the series proves the result:

$$\sum_{\zeta\in\Xi} |\alpha_{\xi\zeta}| \leq C q_\Xi^{3-2m} \frac{e^{2\nu}\left(e^\nu+1\right)}{\left(e^\nu-1\right)^3} = C' q_\Xi^{3-2m}.$$

□

We now establish a bound on the 1-norm of the difference between the collocation matrix and truncated collocation matrix. Again, since both are self-adjoint, the corresponding result for the 2-norm follows automatically.

Lemma 6.2.2. *Let K_Ξ and $\widetilde{K}_\Xi$ be the collocation matrix and truncated collocation matrix, respectively, for κ_m and $\widetilde{\kappa}_m$, respectively, with $m > 5/2$ and truncation parameter $N > 0$. There is a constant C such that, for h_Ξ sufficiently small,*

$$\left\| K_\Xi - \widetilde{K}_\Xi \right\|_1 \leq C h_\Xi^{-3} N^{2-2m}.$$

Proof. By Lemma 6.1.1,

$$\left\| K_\Xi - \widetilde{K}_\Xi \right\|_1 = \max_{\xi\in\Xi} \sum_{\eta\in\Xi} |\kappa_m(\xi,\eta) - \widetilde{\kappa}_m(\xi,\eta)| \leq C\left(\#\Xi\right) N^{2-2m},$$

and $\#\Xi \sim h_\Xi^3$. □

We also need to bound the inverse of the truncated collocation matrix. To do so requires us to assume that the truncation parameter is at least a certain size, given in the following lemma.

Lemma 6.2.3. *Let* $\widetilde{K}_\Xi$ *be the truncated collocation matrix for* $\widetilde{\kappa}_m$*, with* $m > 5/2$ *and truncation parameter* N*. Let* C *be the largest constant from Lemmas 6.2.1 and 6.2.2. Suppose* N *satisfies*

$$N \geq \left(2Ch_\Xi^{-2m}\right)^{\frac{1}{2m-2}}.$$

There exists a constant C' *such that, for* h_Ξ *sufficiently small,*

$$\left\|\widetilde{K}_\Xi^{-1}\right\|_1 \leq C' h_\Xi^{3-2m}.$$

Proof. Note that

$$\widetilde{K}_\Xi = K_\Xi \left(I - K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right),$$

and hence

$$\widetilde{K}_\Xi^{-1} = \left(I - K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right)^{-1} K_\Xi^{-1}.$$

By Lemmas 6.2.1 and 6.2.2, then, along with the condition on N, we have

$$\begin{aligned}\left\|K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right\|_1 &\leq \left\|K_\Xi^{-1}\right\|_1 \left\|K_\Xi - \widetilde{K}_\Xi\right\|_1 \\ &\leq Ch_\Xi^{-2m} N^{2-2m} \leq \frac{1}{2}.\end{aligned}$$

This ensures the convergence of the Neumann series

$$\left(I - K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right)^{-1} = \sum_{n=0}^{\infty} \left(K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right)^n,$$

and we have

$$\left\| \left(I - K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right)^{-1} \right\|_1 \le \sum_{n=0}^{\infty} \left\| K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right\|_1^n$$
$$\le \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 2.$$

Hence by Lemma 6.2.1,

$$\left\| \widetilde{K}_\Xi^{-1} \right\|_1 \le \left\| \left(I - K_\Xi^{-1}\left(K_\Xi - \widetilde{K}_\Xi\right)\right)^{-1} \right\|_1 \left\| K_\Xi^{-1} \right\|_1$$
$$\le C h_\Xi^{3-2m}.$$

□

Corollary 6.2.4. *Adopt the notation and assumptions from Lemma 6.2.3. There is a constant C such that, for h_Ξ sufficiently small,*

$$\|\widetilde{\chi}_\xi\|_{H^m} \le C h_\Xi^{3-2m}.$$

Proof. It is straightforward to show that $\|\widetilde{\kappa}_m(\cdot,\eta)\|_{H^m}$ is bounded for each $\eta \in SO(3)$ by $C = \sqrt{1 + 3\zeta(2m-1)}$, where ζ is the Riemann zeta function. Now,

$$\|\widetilde{\chi}_\xi\|_{H^m} = \left\| \sum_{\eta\in\Xi} \widetilde{\alpha}_{\xi,\eta}\widetilde{\kappa}_m(\cdot,\eta) \right\|_{H^m} \le \sum_{\eta\in\Xi} |\widetilde{\alpha}_{\xi,\eta}| \, \|\widetilde{\kappa}_m(\cdot,\eta)\|_{H^m}$$
$$\le C \sum_{\eta\in\Xi} |\widetilde{\alpha}_{\xi,\eta}| = C \, \|\widetilde{\alpha}_\xi\|_{\ell_1(\Xi)} = C \left\| \widetilde{K}_\Xi^{-1}\delta_\xi \right\|_{\ell_1(\Xi)} \le C \left\| \widetilde{K}_\Xi^{-1} \right\|_1,$$

and the result now follows from Lemma 6.2.3. □

We now turn to a result that is analogous to Theorem 4.1.5. That theorem was used to bound the entries, and then the 1-norm, of the difference between the stiffness matrix and the quadratized stiffness matrix. This result will be used in the next section to bound the entries of the difference between the stiffness matrix and the truncated stiffness matrix as a function of their distance from the main diagonal. That will then be used to get a bound on the 1-norm of that difference.

Theorem 6.2.5. *Let $\{\widetilde{\chi}_\xi\}_{\xi\in\Xi}$ be the truncated Lagrange basis for $\widetilde{\kappa}_m$, with $m > 5/2$. Then $b\widetilde{\chi}_\xi\widetilde{\chi}_\eta \in H^m \cap L_\infty$ and $a\left(\nabla\widetilde{\chi}_\xi, \nabla\widetilde{\chi}_\eta\right) \in H^{m-1}$. Moreover, there is a constant C such that, for h_Ξ sufficiently small,*

$$\|b\widetilde{\chi}_\xi\widetilde{\chi}_\eta\|_{H^m} \leq Ch_\Xi^{3-2m}$$

and

$$\left\|a^\sharp\left(\nabla\widetilde{\chi}_\xi, \nabla\widetilde{\chi}_\eta\right)\right\|_{H^{m-1}} \leq Ch_\Xi^{2-m}.$$

Proof. The proof of this result is identical to the proof of Theorem 4.1.5, except that we use the bound from Corollary 6.2.4 instead of a bump estimate. □

We end this section with results that bound the truncation error for Lagrange functions and their covariant derivatives. They will be instrumental in proving truncation error estimates for the truncated Galerkin approximation.

Lemma 6.2.6. *Let $\{\chi_\xi\}_{\xi\in\Xi}$ and $\{\widetilde{\chi}_\xi\}_{\xi\in\Xi}$ be the Lagrange basis and truncated Lagrange basis, respectively, with $m > 5/2$. Suppose the truncation parameter N satisfies the lower bound in Lemma 6.2.3. There is a constant C' such that, for h_Ξ sufficiently small,*

$$\|\chi_\xi - \widetilde{\chi}_\xi\|_{L_2} \leq C' h_\Xi^{\frac{3}{2}-4m} N^{2-2m}.$$

Proof. By "smuggling in" the "hybrid" function $\sum_{\eta\in\Xi}\alpha_{\xi\eta}\widetilde{\kappa}(\cdot,\eta)$, which pairs the coefficients of the Lagrange functions with the corresponding translates of the truncated kernel, we have

$$\begin{aligned}\|\chi_\xi-\widetilde{\chi}_\xi\|_{L_2} &= \left\|\sum_{\eta\in\Xi}\alpha_{\xi\eta}\kappa_m(\cdot,\eta)-\sum_{\eta\in\Xi}\widetilde{\alpha}_{\xi\eta}\widetilde{\kappa}_m(\cdot,\eta)\right\|_{L_2}\\ &\le \left\|\sum_{\eta\in\Xi}\alpha_{\xi\eta}\left(\kappa_m(\cdot,\eta)-\widetilde{\kappa}_m(\cdot,\eta)\right)\right\|_{L_2}\\ &\qquad + \left\|\sum_{\eta\in\Xi}\left(\alpha_{\xi\eta}-\widetilde{\alpha}_{\xi\eta}\right)\widetilde{\kappa}_m(\cdot,\eta)\right\|_{L_2}.\end{aligned} \tag{6.2.7}$$

We handle the two summands in (6.2.7) separately. For the first, Hölder's inequality for sums gives

$$\begin{aligned}&\left\|\sum_{\eta\in\Xi}\alpha_{\xi\eta}\left(\kappa_m(\cdot,\eta)-\widetilde{\kappa}_m(\cdot,\eta)\right)\right\|_{L_2}^2\\ &=\int_{SO(3)}\left|\sum_{\eta\in\Xi}\alpha_{\xi\eta}\left(\kappa_m(\cdot,\eta)-\widetilde{\kappa}_m(\cdot,\eta)\right)\right|^2 d\mu\\ &\le\int_{SO(3)}\left(\sum_{\eta\in\Xi}|\alpha_{\xi\eta}|\,|\kappa_m(\cdot,\eta)-\widetilde{\kappa}_m(\cdot,\eta)|\right)^2 d\mu\\ &\le\int_{SO(3)}\left(\left(\sum_{\eta\in\Xi}|\alpha_{\xi\eta}|^2\right)^{1/2}\left(\sum_{\zeta\in\Xi}|\kappa_m(\cdot,\zeta)-\widetilde{\kappa}_m(\cdot,\zeta)|^2\right)^{1/2}\right)^2 d\mu\\ &=\sum_{\eta\in\Xi}|\alpha_{\xi\eta}|^2\sum_{\zeta\in\Xi}\|\kappa_m(\cdot,\zeta)-\widetilde{\kappa}_m(\cdot,\zeta)\|_{L_2}^2\end{aligned} \tag{6.2.8}$$

We bound the two factors in (6.2.8) separately. For the first, note that since K_Ξ^{-1} is self-adjoint, $\left\|K_\Xi^{-1}\right\|_2 \le \left\|K_\Xi^{-1}\right\|_1$. Thus, Lemma 6.2.1 gives

$$\sum_{\eta\in\Xi} |\alpha_{\xi\eta}|^2 = \|\alpha_\xi\|^2_{\ell_2(\Xi)} = \left\|K_\Xi^{-1}\delta_\xi\right\|^2_{\ell_2(\Xi)} \le \left\|K_\Xi^{-1}\right\|_2^2 \le \left\|K_\Xi^{-1}\right\|_1^2 \le Ch_\Xi^{6-4m}. \tag{6.2.9}$$

For the second factor in (6.2.8), we have by Lemma 6.1.2 that

$$\sum_{\zeta\in\Xi} \|\kappa_m(\cdot,\zeta) - \widetilde{\kappa}_m(\cdot,\zeta)\|^2_{L_2} \le (\#\Xi)\, CN^{3-4m} \le Cq_\Xi^{-3} N^{4-4m}. \tag{6.2.10}$$

Using $q_\Xi^{-3} = \rho^3 h_\Xi^{-3}$, putting (6.2.9) and (6.2.10) into (6.2.8), and taking square roots, we arrive at

$$\left\|\sum_{\eta\in\Xi} \alpha_{\xi\eta}\left(\kappa_m(\cdot,\eta) - \widetilde{\kappa}_m(\cdot,\eta)\right)\right\|_{L_2} \le Ch_\Xi^{\frac{3}{2}-2m} N^{2-2m}. \tag{6.2.11}$$

We now turn to the second summand in (6.2.7). Similar considerations give

$$\begin{aligned} &\left\|\sum_{\eta\in\Xi} \left(\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}\right)\widetilde{\kappa}_m(\cdot,\eta)\right\|^2_{L_2} \\ &\qquad \le \sum_{\eta\in\Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}|^2 \sum_{\zeta\in\Xi} \|\widetilde{\kappa}_m(\cdot,\zeta)\|^2_{L_2}. \end{aligned} \tag{6.2.12}$$

As before, we bound the two factors in (6.2.12) separately. For the first, we start with

$$\begin{aligned} \sum_{\eta\in\Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}|^2 &= \|\alpha_\xi - \widetilde{\alpha}_\xi\|^2_{\ell_2(\Xi)} \\ &= \left\|\left(K_\Xi^{-1} - \widetilde{K}_\Xi^{-1}\right)\delta_\xi\right\|^2_{\ell_2(\Xi)} \\ &\le \left\|K_\Xi^{-1} - \widetilde{K}_\Xi^{-1}\right\|^2_2 \end{aligned}$$

Both K_Ξ^{-1} and $\widetilde{K}_\Xi^{-1}$ are self-adjoint, and so

$$\left\| K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} \right\|_2 \le \left\| K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} \right\|_1 .$$

Also, we have $K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} = \widetilde{K}_\Xi^{-1} \left(\widetilde{K}_\Xi - K_\Xi \right) K_\Xi^{-1}$. Hence by Lemmas 6.2.1, 6.2.2, and 6.2.3,

$$\begin{aligned} \sum_{\eta \in \Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}|^2 &\le \left\| K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} \right\|_1^2 \\ &\le \left\| \widetilde{K}_\Xi^{-1} \right\|_1^2 \left\| \widetilde{K}_\Xi - K_\Xi \right\|_1^2 \left\| K_\Xi^{-1} \right\|_1^2 \\ &\le C h_\Xi^{6-8m} N^{4-4m}. \end{aligned} \tag{6.2.13}$$

For the second factor in (6.2.12), absorb $\|\widetilde{\kappa}_m(\cdot, \zeta)\|_{L_2}^2$, which depends only on m, into the constant. This gives us

$$\sum_{\zeta \in \Xi} \|\widetilde{\kappa}_m(\cdot, \zeta)\|_{L_2}^2 \le C\,(\#\Xi) \le C q_\Xi^{-3}. \tag{6.2.14}$$

Again using $q_\Xi^{-3} = \rho^3 h_\Xi^{-3}$, putting (6.2.13) and (6.2.14) into (6.2.12), and taking square roots, we arrive at

$$\left\| \sum_{\eta \in \Xi} (\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta})\, \widetilde{\kappa}_m(\cdot, \zeta) \right\|_{L_2} \le C h_\Xi^{\frac{3}{2} - 4m} N^{2-2m}. \tag{6.2.15}$$

Finally, then, putting (6.2.11) and (6.2.15) into (6.2.7) yields the result. □

An almost identical proof, where we use Lemma 6.1.3 instead of 6.1.2, yields the following.

Lemma 6.2.16. *Let $\{\chi_\xi\}_{\xi \in \Xi}$ and $\{\widetilde{\chi}_\xi\}_{\xi \in \Xi}$ be the Lagrange basis and truncated Lagrange basis, respectively, for κ_m and $\widetilde{\kappa}_m$, respectively, with $m > 5/2$. Suppose the truncation parameter N satisfies the lower bound in*

Lemma 6.2.3. There is a constant C' such that, for h_Ξ sufficiently small,

$$\|\nabla(\chi_\xi - \widetilde{\chi}_\xi)\|_{L_2} \le C' h_\Xi^{-4m} N^{3-2m}$$

We also need an analogous result for the Sobolev norm of the difference between the Lagrange function and the truncated Lagrange function.

Lemma 6.2.17. *Let $\{\chi_\xi\}_{\xi\in\Xi}$ and $\{\widetilde{\chi}_\xi\}_{\xi\in\Xi}$ be the Lagrange basis and truncated Lagrange basis, respectively, with $m > 5/2$. Suppose the truncation parameter N satisfies the lower bound in Lemma 6.2.3. There is a constant C' such that, for h_Ξ sufficiently small,*

$$\|\chi_\xi - \widetilde{\chi}_\xi\|_{H^m} \le C' h_\Xi^{-4m} N^{2-2m}.$$

Proof. By "smuggling in" the "hybrid" function $\sum_{\eta\in\Xi} \alpha_{\xi\eta}\widetilde{\kappa}(\cdot,\eta)$, which pairs the coefficients of the Lagrange functions with the corresponding translates of the truncated kernel, we have

$$\begin{aligned}\|\chi_\xi - \widetilde{\chi}_\xi\|_{H^m} &= \left\| \sum_{\eta\in\Xi} \alpha_{\xi\eta}\kappa_m(\cdot,\eta) - \sum_{\eta\in\Xi} \widetilde{\alpha}_{\xi\eta}\widetilde{\kappa}_m(\cdot,\eta) \right\|_{H^m} \\ &\le \left\| \sum_{\eta\in\Xi} \alpha_{\xi\eta}\left(\kappa_m(\cdot,\eta) - \widetilde{\kappa}_m(\cdot,\eta)\right) \right\|_{H^m} \\ &\qquad + \left\| \sum_{\eta\in\Xi} \left(\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}\right)\widetilde{\kappa}_m(\cdot,\eta) \right\|_{H^m}. \end{aligned} \tag{6.2.18}$$

We handle the sums in (6.2.18) separately. For the first, we've already seen

$\sum_{\eta\in\Xi} |\alpha_{\xi\eta}| \le Ch_\Xi^{3-2m}$, and so Lemma 6.1.2 gives

$$\left\| \sum_{\eta\in\Xi} (\kappa_m(\cdot,\eta) - \widetilde{\kappa}_m(\cdot,\eta)) \right\|_{H^m} \le \sum_{\eta\in\Xi} |\alpha_{\xi\eta}| \, \|\kappa_m(\cdot,\eta) - \widetilde{\kappa}_m(\cdot,\eta)\|_{H^m} \le Ch_\Xi^{3-2m} N^{2-2m}. \tag{6.2.19}$$

For the second, start with

$$\left\| \sum_{\eta\in\Xi} (\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}) \, \widetilde{\kappa}_m(\cdot,\eta) \right\|_{H^m} \le \sum_{\eta\in\Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}| \, \|\widetilde{\kappa}_m(\cdot,\eta)\|_{H^m} .$$

Now, $\|\widetilde{\kappa}_m(\cdot,\eta)\|_{H^m}$ is bounded by a constant depending only on m. We also have

$$\begin{aligned} \sum_{\eta\in\Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}| &= \|\alpha_\xi - \widetilde{\alpha}_\xi\|_{\ell_1(\Xi)} \\ &= \left\| \left(K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} \right) \delta_\xi \right\|_{\ell_1(\Xi)} \\ &\le \left\| K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} \right\|_1 . \end{aligned}$$

Now, $K_\Xi^{-1} - \widetilde{K}_\Xi^{-1} = \widetilde{K}_\Xi^{-1} \left(\widetilde{K}_\Xi - K_\Xi \right) K_\Xi^{-1}$, and so by Lemmas 6.2.1, 6.2.2, and 6.2.3,

$$\begin{aligned} \sum_{\eta\in\Xi} |\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}| &\le \left\| \widetilde{K}_\Xi^{-1} \right\|_1 \left\| \widetilde{K}_\Xi - K_\Xi \right\|_1 \left\| K_\Xi^{-1} \right\|_1 \\ &\le Ch_\Xi^{3-4m} N^{2-2m} \end{aligned}$$

Therefore

$$\left\| \sum_{\eta\in\Xi} (\alpha_{\xi\eta} - \widetilde{\alpha}_{\xi\eta}) \, \widetilde{\kappa}_m(\cdot,\zeta) \right\|_{H^m} \le Ch_\Xi^{-4m} N^{2-2m} \tag{6.2.20}$$

The result now follows by putting (6.2.19) and (6.2.20) into (6.2.7). □

6.3 The Truncated Stiffness Matrix

The *truncated stiffness matrix* in the truncated Lagrange basis, $\widetilde{B}_\Xi$, has entries $\widetilde{B}_{\xi\eta} = \langle \widetilde{\chi}_\xi, \widetilde{\chi}_\eta \rangle_{a,b}$. We first establish a bound on the entries of the difference between the stiffness matrix and the truncated stiffness matrix, which allows us to get a bound on the 1- and 2-norms of that difference.

Theorem 6.3.1. *Let B_Ξ and $\widetilde{B}_\Xi$ be the stiffness matrix and truncated stiffness matrix, respectively, for κ_m and $\widetilde{\kappa}_m$, respectively, with $m > 5/2$, and centers Ξ. Assume the truncation parameter N satisfies the lower bound from Lemma 6.2.3. There is a constant C such that, for h_Ξ sufficiently small,*

$$\left| B_{\xi\eta} - \widetilde{B}_{\xi\eta} \right| \leq C h_\Xi^{-4m} N^{2-2m}.$$

Proof. First, we "smuggle in" $\chi_\xi \widetilde{\chi}_\eta$ and apply Hölder's inequality:

$$\begin{aligned} \left| \int_{SO(3)} b \left(\chi_\xi \chi_\eta - \widetilde{\chi}_\xi \widetilde{\chi}_\eta \right) d\mu \right| &\leq b_2 \int_{SO(3)} \left(\left| \chi_\xi \left(\chi_\eta - \widetilde{\chi}_\eta \right) \right| \right. \\ &\qquad \left. + \left| \left(\chi_\xi - \widetilde{\chi}_\xi \right) \widetilde{\chi}_\eta \right| \right) d\mu \\ &\leq b_2 \Big(\left\| \chi_\xi \right\|_{L_2} \left\| \chi_\eta - \widetilde{\chi}_\eta \right\|_{L_2} \\ &\qquad + \left\| \chi_\xi - \widetilde{\chi}_\xi \right\|_{L_2} \left\| \widetilde{\chi}_\eta \right\|_{L_2} \Big). \end{aligned}$$

By Lemma 6.2.6, our bump estimate, and Corollary 6.2.4, then,

$$\left| \int_{SO(3)} b \left(\chi_\xi \chi_\eta - \widetilde{\chi}_\xi \widetilde{\chi}_\eta \right) d\mu \right| \leq C h_\Xi^{\frac{9}{2} - 4m} N^{2-2m} \tag{6.3.2}$$

We proceed now exactly as we did in the second part of the proof of Lemma 4.1.5. Cover $SO(3)$ with finitely many Ω_k with $\Omega_k \subseteq b(q_k, r_{SO(3)})$, and let $U_k = \mathrm{Exp}_k^{-1}(\Omega_k)$. Let $\{\tau_k\}_{k \in \{1,\ldots,K\}}$ be a partition of unity subordinate to that finite cover. The exact same arguments give

$$\begin{aligned}
&\left|\int_{SO(3)} a^{\sharp}\left(\nabla\left(\chi_\xi-\widetilde{\chi}_\xi\right),\nabla\left(\chi_\eta-\widetilde{\chi}_\eta\right)\right)d\mu\right|\\
&\leq \int_{SO(3)}\left|a^{\sharp}\left(\nabla\left(\chi_\xi-\widetilde{\chi}_\xi\right),\nabla\left(\chi_\eta-\widetilde{\chi}_\eta\right)\right)\right|d\mu\\
&=\left\|a^{\sharp}\left(\nabla\left(\chi_\xi-\widetilde{\chi}_\xi\right),\nabla\left(\chi_\eta-\widetilde{\chi}_\eta\right)\right)\right\|_{L_2}\\
&\leq \sum_{k=1}^{K}\sum_{i,j}\left\|\tau_k a^{ij}\frac{\partial\left(\chi_\xi-\widetilde{\chi}_\xi\right)}{\partial x_i}\frac{\partial\left(\chi_\eta-\widetilde{\chi}_\eta\right)}{\partial x_j}\right\|_{L_2}.
\end{aligned}$$

Also the exact same arguments give (4.1.8) with χ_ξ and χ_η replaced with $\chi_\xi-\widetilde{\chi}$ and $\chi_\eta-\widetilde{\chi}_\eta$, respectively. It thus becomes a matter of bounding $\left\|\sigma\frac{\partial(\chi_\xi-\widetilde{\chi}_\xi)}{\partial x_i}\right\|_{L_2}$, where σ is the same cutoff function. Again using the exact same arguments, we obtain

$$\left\|\sigma\frac{\partial\left(\chi_\xi-\widetilde{\chi}_\xi\right)}{\partial x_i}\right\|_{L_2}\leq C\left\|\chi_\xi-\widetilde{\chi}_\xi\right\|_{H^1}.$$

The Zeros Lemma applies to $\chi_\xi-\widetilde{\chi}_\xi$, since it vanishes on Ξ. That and Lemma 6.2.6 give

$$\left\|\sigma\frac{\partial\left(\chi_\xi-\widetilde{\chi}_\xi\right)}{\partial x_i}\right\|_{L_2}\leq Ch_\Xi^{m-1}\left\|\chi_\xi-\widetilde{\chi}_\xi\right\|_{H^m}\leq Ch_\Xi^{-3m-1}N^{2-2m}.$$

Putting this all together gives

$$\left|\int_{SO(3)}\left(a^{\sharp}\left(\nabla\chi_{xi},\nabla\chi_\eta\right)-a^{\sharp}\left(\nabla\widetilde{\chi}_\xi,\nabla\widetilde{\chi}_\eta\right)\right)d\mu\right|\leq Ch_\Xi^{-3m-1}N^{2-2m},\quad(6.3.3)$$

and the result follows from (6.3.2), (6.3.3), and the triangle inequality. □

Corollary 6.3.4. *Adopt the notation and assumptions of Theorem 6.3.1.*

There is a constant C such that, for h_Ξ sufficiently small,

$$\left\| B_\Xi - \widetilde{B}_\Xi \right\|_2 \leq C h_\Xi^{-3-4m} N^{2-2m}.$$

Proof. Since B_Ξ and $\widetilde{B}_\Xi$ are self-adjoint,

$$\left\| B_\Xi - \widetilde{B}_\Xi \right\|_2 \leq \left\| B_\Xi - \widetilde{B}_\Xi \right\|_1 = \max_{\xi\in\Xi} \sum_{\eta\in\Xi} \left| B_{\xi\eta} - \widetilde{B}_{\xi\eta} \right|.$$

For any $\xi \in \Xi$, we have by Theorem 6.3.1 that

$$\sum_{\eta\in\Xi} \left| B_{\xi\eta} - \widetilde{B}_{\xi\eta} \right| \leq (\#\Xi) \cdot C h_\Xi^{-4m} N^{2-2m},$$

and the result follows from $\#\Xi \sim q_\Xi^{-3} = \rho^3 h_\Xi^{-3}$. □

We need a result similar to the one in Lemma 4.1.25. There we needed to bound the 1- and 2- norms of the inverse of the quadratized stiffness matrix. Here we need to bound the 2-norm of the inverse of the truncated stiffness matrix. This result requires a lower bound on the truncation parameter, given in the statement.

Theorem 6.3.5. *Adopt the notation and assumptions of Theorem 6.3.1. Let C_0 be the largest constant from Lemma 6.2.1, Lemma 6.2.2, and Corollary 6.3.4. Assume the truncation parameter N satisfies*

$$N \geq \left(2C_0 h_\Xi^{-3-4m}\right)^{\frac{1}{2m-3}}.$$

There is a constant C' such that, for h_Ξ sufficiently small,

$$\left\| \widetilde{B}_\Xi^{-1} \right\|_2 \leq C' h_\Xi^{-3}.$$

Proof. Note that

$$\widetilde{B}_\Xi = B_\Xi \left(I - B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right),$$

and hence

$$\widetilde{B}_\Xi^{-1} = \left(I - B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right)^{-1} B_\Xi^{-1}.$$

By Theorem 3.4.4 and Corollary 6.3.4,

$$\begin{aligned} \left\| B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right\|_2 &\le \|B_\Xi^{-1}\|_2 \left\| B_\Xi - \widetilde{B}_\Xi \right\|_2 \\ &\le C h_\Xi^{-3-4m} N^{3-2m}. \end{aligned}$$

The assumption on N ensures that

$$\left\| B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right\|_2 \le \frac{1}{2},$$

which in turn ensures the convergence of the Neumann series

$$\left(I - B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right)^{-1} = \sum_{n=0}^{\infty} \left(B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right)^n.$$

Therefore

$$\begin{aligned} \left\| \left(I - B_\Xi^{-1} \left(B_{Xi} - \widetilde{B}_\Xi \right) \right)^{-1} \right\|_2 &\le \sum_{n=0}^{\infty} \left\| B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right\|_2^n \\ &\le \sum_{n=0}^{\infty} \left(\frac{1}{2} \right)^n = 2, \end{aligned}$$

and hence

$$\left\| \widetilde{B}_\Xi^{-1} \right\|_2 \le \left\| \left(I - B_\Xi^{-1} \left(B_\Xi - \widetilde{B}_\Xi \right) \right)^{-1} \right\|_2 \|B_\Xi^{-1}\|_2 \le C \|B_\Xi^{-1}\|_2 \le C h_\Xi^{-3}.$$

$\square$

6.4 Truncated Galerkin Approximation

The *truncated Galerkin approximation* is

$$\widetilde{u}_\Xi = \sum_{\xi \in \Xi} \widetilde{\omega}_\xi \widetilde{\chi}_\xi,$$

where the coefficients $\widetilde{\gamma} = \{\widetilde{\gamma}_\xi\}_{\xi\in\Xi}$ are obtained by solving the linear system $\widetilde{B}_\Xi \gamma = \widetilde{\omega}$, with $\widetilde{\omega} = \{\widetilde{\omega}_\xi\}_{\xi\in\Xi}$ the vector with entries

$$\widetilde{\omega}_\xi = \int_{SO(3)} f \widetilde{\chi}_\xi d\mu.$$

Armed with all the results we've established thus far, we're ready to prove our truncated Galerkin approximation error estimate right away. As usual, it will be immediately followed by a result that estimates the error between the weak solution and the truncated Galerkin approximation - it is just an application of the triangle inequality.

Theorem 6.4.1. (Truncated Galerkin Approximation Error Estimate) *Let u_Ξ and $\widetilde{u}_\Xi$ be the Galerkin approximation and truncated Galerkin approximation, respectively, for κ_m and $\widetilde{\kappa}_m$, respectively, with $m > 5/2$, centers Ξ. (If κ_m is conditionally positive definite with respect to Π_L, assume Ξ is Π_L-unisolvent.) Assume the truncation parameter N satisfies the lower bound in Theorem 6.3.5. Let $f \in H^s$, where $s = m-1$ if $5/2 < m \le 7/2$ and let $s = m-2$ if $m > 7/2$. There is a constant C such that, for h_Ξ sufficiently small,*

$$\|u_\Xi - \widetilde{u}_\Xi\|_{L_2} \le C h_\Xi^{-\frac{15}{2}-4m} N^{2-2m} \|f\|_{H^s}.$$

Remark 6.4.2. The reason for the different cases for s is the same as the one given in Remark 4.2.2, we just have $d = 3$ now.

Proof. We start by "smuggling in" $\sum_{\xi \in \Xi} \gamma_\xi \widetilde{\chi}_\xi$ and applying the triangle inequality:

$$\|u_\Xi - \widetilde{u}_\Xi\|_{L_2} \leq \left\| \sum_{\xi \in \Xi} \gamma_\xi \left(\chi_\xi - \widetilde{\chi}_\xi\right) \right\|_{L_2} + \left\| \sum_{\xi \in \Xi} \left(\gamma - \widetilde{\gamma}\right) \widetilde{\chi}_\xi \right\|_{L_2}. \tag{6.4.3}$$

We handle the summands in (6.4.3) separately. For the first, we have that, as in the proof of Lemma 6.2.6, Hölder's inequality for sums gives

$$\left\| \sum_{\xi \in \Xi} \gamma_\xi \left(\chi_\xi - \widetilde{\chi}_\xi\right) \right\|_{L_2}^2 \leq \sum_{\eta \in \Xi} |\gamma_\eta|^2 \sum_{\zeta \in \Xi} \|\chi_\zeta - \widetilde{\chi}_\zeta\|_{L_2}^2. \tag{6.4.4}$$

Similarly, for the second summand in (6.4.3),

$$\left\| \sum_{\xi \in \Xi} \left(\gamma_\xi - \widetilde{\gamma}_\xi\right) \widetilde{\chi}_\xi \right\|_{L_2}^2 \leq \sum_{\eta \in \Xi} |\gamma_\xi - \widetilde{\gamma}_\xi|^2 \sum_{\zeta \in \Xi} \|\widetilde{\chi}_\zeta\|_{L_2}^2. \tag{6.4.5}$$

We handle the two factors in (6.4.4) separately. For the first, we have by the L_2-stability of the Lagrange basis that

$$\begin{aligned}
\sum_{\eta \in \Xi} |\gamma_\eta|^2 = \|\gamma\|_{\ell_2(\Xi)}^2 &\leq C h_\Xi^{-3} \left\| \sum_{\xi \in \Xi} \gamma_\xi \chi_\xi \right\|_{L_2}^2 \\
&= C h_\Xi^{-3} \|u_\Xi\|_{L_2}^2 \leq C h_\Xi^{-3} \left(\|u_\Xi - u\|_{L_2} + \|u\|_{L_2}\right)^2.
\end{aligned}$$

We've already seen $\|u - u_\Xi\|_{L_2} \leq C h_\Xi^{m-1} \|f\|_{H^s}$, and $\|u\|_{L_2} \leq \|u\|_{H_2} \leq$

$C\|f\|_{L_2} \le C\|f\|_{H^s}$, and hence

$$\sum_{\eta\in\Xi} |\gamma_\xi|^2 \le C h_\Xi^{-3} \left(h_\Xi^{m-1}+1\right)^2 \|f\|_{H^s}^2 \le C h_\Xi^{-3}\|f\|_{H^s}^2. \tag{6.4.6}$$

For the second factor in (6.4.4), we have by Theorem 6.2.6 that

$$\begin{aligned}\sum_{\zeta\in\Xi} \|\chi_\xi - \widetilde{\chi}_\xi\|_{L_2}^2 &\le (\#\Xi)\cdot C h_\Xi^{3-8m} N^{4-4m} \\ &\le C h_\Xi^{-8m} N^{4-4m}.\end{aligned} \tag{6.4.7}$$

Putting (6.4.6) and (6.4.7) into (6.4.4) and taking square roots, we arrive at

$$\left\|\sum_{\xi\in\Xi} \gamma_\xi \left(\chi_\xi - \widetilde{\chi}_\xi\right)\right\|_{L_2} \le C h_\Xi^{-\frac{3}{2}-4m} N^{2-2m} \|f\|_{H^s}. \tag{6.4.8}$$

We now tackle the product in (6.4.5) by handling each factor separately. For the second factor, we have $\|\widetilde{\chi}_\xi\|_{L_2}^2 \le C$, so

$$\sum_{\zeta\in\Xi} \|\widetilde{\chi}_\xi\|_{L_2}^2 \le (\#\Xi)\cdot C \le C h_\Xi^{-3}. \tag{6.4.9}$$

For the first factor in (6.4.5), we start by "smuggling in" $\widetilde{B}_\Xi^{-1}\omega$ to obtain

$$\begin{aligned}\sum_{\eta\in\Xi} |\gamma_\eta - \widetilde{\gamma}_\eta|^2 = \|\gamma - \widetilde{\gamma}\|_{\ell_2(\Xi)}^2 &= \left\|B_\Xi^{-1}\omega - \widetilde{B}_\Xi^{-1}\widetilde{\omega}\right\|_{\ell_2(\Xi)}^2 \\ &\le \left(\left\|\left(B_\Xi^{-1} - \widetilde{B}_\Xi^{-1}\right)\omega\right\|_{\ell_2(\Xi)} + \left\|\widetilde{B}_\Xi^{-1}\left(\omega - \widetilde{\omega}\right)\right\|_{\ell_2(\Xi)}\right)^2 \\ &\le \left(\left\|\widetilde{B}_\Xi^{-1}\right\|_2 \left\|\widetilde{B}_\Xi - B_\Xi\right\|_2 \left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} \right. \\ &\qquad \left. + \left\|\widetilde{B}_\Xi^{-1}\right\|_2 \|\omega - \widetilde{\omega}\|_{\ell_2(\Xi)}\right)^2.\end{aligned}$$

As we saw in the proof of Theorem 4.2.1, $\left\|B_\Xi^{-1}\omega\right\|_{\ell_2(\Xi)} = \|\gamma\|_{\ell_2(\Xi)} \le$

$Ch_\Xi^{-3/2}\|f\|_{H^s}$, and Theorems 6.3.4 and 6.3.5 give

$$\sum_{\eta\in\Xi}|\gamma_\eta-\widetilde{\gamma}_\eta|^2\le\left(Ch_\Xi^{-6-4m}N^{2-2m}\|f\|_{H^s}+Ch_\Xi^{-3}\|\omega-\widetilde{\omega}\|_{\ell_2(\Xi)}\right)^2. \tag{6.4.10}$$

We need to deal with the second term in the parentheses in (6.4.10). We start by using Hölder's inequality to get

$$\begin{aligned}\|\omega-\widetilde{\omega}\|_{\ell_2(\Xi)}^2 &= \sum_{\xi\in\Xi}|\omega_\xi-\widetilde{\omega}_\xi|^2\\ &= \sum_{\xi\in\Xi}\left|\int_{SO(3)}(\chi_\xi-\widetilde{\chi}_\xi)\,f\;d\mu\right|^2\\ &= \|f\|_{L_2}^2\sum_{\xi\in\Xi}\|\chi_\xi-\widetilde{\chi}_\xi\|_{L_2}^2.\end{aligned}$$

Then (6.4.7) and $\|f\|_{L_2}\le\|f\|_{H^s}$ give

$$\|\omega-\widetilde{\omega}\|_{\ell_2(\Xi)}^2\le Ch_\Xi^{-8m}N^{4-4m}\|f\|_{H^m}^2. \tag{6.4.11}$$

Putting the square root of (6.4.11) into (6.4.10) gives

$$\sum_{\eta\in\Xi}|\gamma_\eta-\widetilde{\gamma}_\eta|^2\le Ch_\Xi^{-12-8m}N^{6-4m}\|f\|_{H^s}^2, \tag{6.4.12}$$

and putting (6.4.12) and (6.4.9) into (6.4.5) and taking square roots, we arrive at

$$\left\|\sum_{\xi\in\Xi}(\gamma_\xi-\widetilde{\gamma}_\xi)\,\widetilde{\chi}_\xi\right\|_{L_2}\le Ch_\Xi^{-\frac{15}{2}-4m}N^{2-2m}\|f\|_{H^s}. \tag{6.4.13}$$

Finally, putting (6.4.8) and (6.4.13) into (6.4.3) yields the result. □

Corollary 6.4.14. *Adopt the notation and assumptions of Theorem 6.4.1.*

There is a constant C such that, for h_Ξ sufficiently small,

$$\|u - \widetilde{u}_\Xi\|_{L_2} \le C \left(h_\Xi^{m-1} + h_\Xi^{-\frac{15}{2}-4m} N^{2-2m} \right) \|f\|_{H^s}$$

Remark 6.4.15. This brings us to another main point: we can get the same result with truncation as we can without by choosing the truncation parameter N large enough. We already have the lower bound from Theorem 6.5.3. If we also have $N \ge h_\Xi^{-\frac{5}{2}-\frac{23}{4m-4}}$, then the truncation error, $h_\Xi^{-\frac{15}{2}-4m} N^{2-2m}$, will be at most the Galerkin approximation error, h_Ξ^{m-1}.

6.5 The Quadratized Truncated Stiffness Matrix

The *quadratized truncated stiffness matrix*, $\widetilde{B}_\Xi^\Lambda = \left\{ \widetilde{B}_{\xi\eta}^\Lambda \right\}$ is obtained by replacing each entry in the truncated stiffness matrix with the quadrature estimate

$$\widetilde{B}_{\xi\eta}^\Lambda = Q^\Lambda \left(a^\sharp \left(\nabla \widetilde{\chi}_\xi, \nabla \widetilde{\chi}_\eta \right) + b \widetilde{\chi}_\xi \widetilde{\chi}_\eta \right).$$

Our first order of business in this section is to bound the entries of the difference between the truncated stiffness matrix and the quadratized truncated stiffness matrix. It will be immediately followed by a bound on the 2-norm of that difference.

Theorem 6.5.1. *Let $\widetilde{B}_\Xi$ and $\widetilde{B}_\Xi^\Lambda$ be the truncated stiffness matrix and quadratized stiffness matrix, respectively, for $\widetilde{\kappa}_m$ with $m > 5/2$, where Ξ is the set of centers used for Galerkin approximation and Λ is the set of centers used for quadrature. (If κ_m is conditionally positive definite with respect to Π_L, assume Ξ is Π_L-unisolvent.) Assume $\#\Xi < \#\Lambda$ and $h_\Lambda < h_\Xi$. Suppose the truncation parameter N satisfies the lower bound in Theorem*

6.3.5. There is a constant C such that, for h_Ξ sufficiently small,

$$\left|\widetilde{B}_{\xi\eta} - \widetilde{B}^{\Lambda}_{\xi\eta}\right| \leq C h_\Xi^{2-2m} h_\Lambda^{m-1}$$

for all $\xi, \eta \in \Xi$.

Proof. By Theorems 3.1.1 and 6.2.5,

$$\begin{aligned}\left|\int_{SO(3)} b\widetilde{\chi}_\xi \widetilde{\chi}_\eta d\mu - Q^\Lambda\left(b\widetilde{\chi}_\xi\widetilde{\chi}_\eta\right)\right| &\leq C h_\Lambda^m \left\|b\widetilde{\chi}_\xi\widetilde{\chi}_\eta\right\|_{H^m} \\ &\leq C h_\Lambda^m h_\Xi^{3-2m}\end{aligned}$$

and

$$\begin{aligned}&\left|\int_{SO(3)} a^\sharp\left(\nabla\widetilde{\chi}_\xi, \nabla\widetilde{\chi}_\eta\right) d\mu - Q^\Lambda\left(a^\sharp\left(\nabla\widetilde{\chi}_\xi, \nabla\widetilde{\chi}_\eta\right)\right)\right| \\ &\leq C h_\Lambda^{m-1}\left\|a^\sharp\left(\nabla\widetilde{\chi}_\xi, \nabla\widetilde{\chi}_\eta\right)\right\|_{H^{m-1}} \\ &\leq C h_\Lambda^{m-1} h_\Xi^{2-2m}.\end{aligned}$$

□

Corollary 6.5.2. *Adopt the notation and assumptions from Theorem 6.5.1. There is a constant C such that, for h_Ξ sufficiently small,*

$$\left\|\widetilde{B}_\Xi - \widetilde{B}^\Lambda_\Xi\right\|_2 \leq C h_\Xi^{-1-2m} h_\Lambda^{m-1}.$$

Proof. This follows the self-adjointness of $\widetilde{B}_\Xi$ and $\widetilde{B}^\Lambda_\Xi$, Theorem 6.5.1, and $\#\Xi \sim q_\Xi^{-3} = \rho^3 h_\Xi^{-3}$:

$$\begin{aligned}\left\|\widetilde{B}_\Xi - \widetilde{B}^\Lambda_\Xi\right\|_2 \leq \left\|\widetilde{B}_\Xi - \widetilde{B}^\Lambda_\Xi\right\|_1 &= \max_{\xi\in\Xi}\sum_{\eta\in\Xi}\left|\widetilde{B}_{\xi\eta} - \widetilde{B}^\Lambda_{\xi\eta}\right| \\ &\leq (\#\Xi)\cdot C h_\Xi^{2-2m} h_\Lambda^{m-1}.\end{aligned}$$

□

We end this section with a result regarding the inverse of the quadratized truncated stiffness matrix. The proof is virtually identical to that of Theorem 6.3.5. As such, we omit it.

Theorem 6.5.3. *Adopt the notation and assumptions from Theorem 6.5.1. Let C_0 be the largest constant from Theorem 6.3.5 and Corollary 6.5.2. Assume*

$$h_\Lambda \le C_0^{\frac{1}{m-1}} h_\Xi^{2+\frac{6}{m-1}} 2^{\frac{1}{1-m}}.$$

There exists a constant C such that, for h_Ξ sufficiently small,

$$\left\| \left(\widetilde{B}_\Xi^\Lambda \right)^{-1} \right\|_2 \le C' h_\Xi^{-3}.$$

6.6 Quadratized Truncated Approximation

The *quadratized truncated Galerkin approximation* is

$$\widetilde{u}_\Xi^\Lambda = \sum_{\xi \in \Xi} \widetilde{\gamma}_\xi^\Lambda \widetilde{\chi}_\xi,$$

where the coefficients $\widetilde{\gamma}^\Lambda = \left\{ \widetilde{\gamma}_\xi^\Lambda \right\}_{\xi \in \Xi}$ are obtained by solving the linear system $\widetilde{B}_\Xi^\Lambda \widetilde{\gamma}^\Lambda = \widetilde{\omega}^\Lambda$, with $\widetilde{\omega}^\Lambda = \left\{ \widetilde{\omega}_\xi^\Lambda \right\}_{\xi \in \Xi}$ the vector with entries

$$\widetilde{\omega}_\xi^\Lambda = Q^\Lambda \left(\widetilde{\chi}_\xi f \right).$$

Armed with the results we have established thus far, we are ready to prove our quadratized truncated Galerkin approximation error estimate right away. It will be immediately followed by an error estimate between the weak solution and the quadratized truncated Galerkin approximation - it is just repeated applications of the triangle inequality.

Theorem 6.6.1. (Quadratized Truncated Galerkin Approximation Error Estimate) *Let* $\widetilde{u}_\Xi$ *and* $\widetilde{u}_\Xi^\Lambda$ *be the truncated Galerkin approximation and the quadratized truncated Galerkin approximation, respectively, for* $\widetilde{\kappa}_m$ *with* $m > 5/2$*, where* Ξ *is the set of centers used for Galerkin approximation and* Λ *is the set of centers used for quadrature. Assume the truncation parameter* N *satisfies the the lower bound from Theorem 6.3.5 and* h_Λ *satisfies both the upper bounds from Lemmas 4.1.25 and 6.5.3. Let* $f \in H^s$*, where* $s = m - 1$ *if* $5/2 < m \le 7/5$ *and* $s = m - 2$ *if* $m > 7/2$*. There is a constant* C *such that, for* h_Ξ *sufficiently small,*

$$\left\|\widetilde{u}_\Xi - \widetilde{u}_\Xi^\Lambda\right\|_{L_2} \le C h_\Xi^{-7-2m} h_\Lambda^{m-1} \|f\|_{H^s}.$$

Proof. As in the proof of Theorem 6.4.1, we use Hölder's inequality for sums to obtain

$$\left\|\widetilde{u}_\Xi - \widetilde{u}_\Xi^\Lambda\right\|_{L_2}^2 \le \sum_{\eta\in\Xi} \left|\widetilde{\gamma}_\eta - \widetilde{\gamma}_\eta^\Lambda\right|^2 \sum_{\zeta\in\Xi} \|\widetilde{\chi}_\zeta\|_{L_2}^2 . \tag{6.6.2}$$

We deal with the factors in (6.6.2) separately. For the first, we start by "smuggling in" $\left(\widetilde{B}_\Xi^\Lambda\right)^{-1} \widetilde{\omega}$:

$$\begin{aligned}
\sum_{\eta\in\Xi} \left|\widetilde{\gamma}_\eta - \widetilde{\gamma}_\eta^\Lambda\right|^2 &= \left\|\widetilde{\gamma} - \widetilde{\gamma}^\Lambda\right\|_{\ell_2(\Xi)}^2 \\
&= \left\|\widetilde{B}_\Xi^{-1}\widetilde{\omega} - \left(\widetilde{B}_\Xi^\Lambda\right)^{-1} \widetilde{\omega}^\Lambda\right\|_{\ell_2(\Xi)}^2 \\
&\le \left(\left\|\left(\widetilde{B}_\Xi^{-1} - \left(\widetilde{B}_\Xi^\Lambda\right)^{-1}\right)\widetilde{\omega}\right\|_{\ell_2(\Xi)}\right. \\
&\qquad \left. + \left\|\left(\widetilde{B}_\Xi^\Lambda\right)^{-1} \left(\widetilde{\omega} - \widetilde{\omega}^\Lambda\right)\right\|_{\ell_2(\Xi)}\right)^2
\end{aligned}$$

Noting that $\widetilde{B}_{\Xi}^{-1} - \left(\widetilde{B}_{\Xi}^{\Lambda}\right)^{-1} = \left(\widetilde{B}_{\Xi}^{\Lambda}\right)^{-1}\left(\widetilde{B}_{\Xi} - \widetilde{B}_{\Xi}^{\Lambda}\right)\widetilde{B}_{\Xi}^{-1}$, we have

$$\sum_{\eta\in\Xi}\left|\widetilde{\gamma}_{\eta} - \widetilde{\gamma}_{\eta}^{\Lambda}\right|^2 \leq \Bigg(\left\|\left(\widetilde{B}_{\Xi}^{\Lambda}\right)^{-1}\right\|_2 \left\|\widetilde{B}_{\Xi} - \widetilde{B}_{\Xi}^{\Lambda}\right\|_2 \left\|\widetilde{B}_{\Xi}^{-1}\widetilde{\omega}\right\|_{\ell_2(\Xi)} + \left\|\left(\widetilde{B}_{\Xi}^{\Lambda}\right)^{-1}\right\|_2 \left\|\widetilde{\omega} - \widetilde{\omega}^{\Lambda}\right\|_{\ell_2(\Xi)}\Bigg)^2$$

Using Theorem 6.5.3,

$$\sum_{\eta\in\Xi}\left|\widetilde{\gamma}_{\eta} - \widetilde{\gamma}_{\eta}^{\Lambda}\right|^2 \leq Ch_{\Xi}^{-6}\left(\left\|\widetilde{B}_{\Xi} - \widetilde{B}_{\Xi}^{\Lambda}\right\|_2 \left\|\widetilde{B}_{\Xi}^{-1}\widetilde{\omega}\right\|_{\ell_2(\Xi)} + \left\|\widetilde{\omega} - \widetilde{\omega}^{\Lambda}\right\|_{\ell_2(\Xi)}\right)^2. \tag{6.6.3}$$

We deal with the terms in the parentheses in (6.6.3) separately. For the second, we have by Theorem 3.1.1 with $\tau = m$ and $\sigma = s$, and 2.7.5,

$$\begin{aligned}
\left\|\widetilde{\omega} - \widetilde{\omega}^{\Lambda}\right\|_{\ell_2(\Xi)}^2 &= \sum_{\xi\in\Xi}\left(\int_{SO(3)} f\widetilde{\chi}_{\xi}\,d\mu - Q^{\Lambda}\left(f\widetilde{\chi}_{\xi}\right)\right)^2 \\
&\leq Ch_{\Lambda}^{2m}\sum_{\xi\in\Xi}\|f\widetilde{\chi}_{\xi}\|_{H^s}^2 \\
&\leq Ch_{\Lambda}^{2m}\sum_{\xi\in\Xi}\left(\|f\|_{H^s}\|\widetilde{\chi}_{\xi}\|_{L_\infty} + \|f\|_{L_\infty}\|\widetilde{\chi}_{\xi}\|_{H^s}\right)^2.
\end{aligned}$$

As we have already seen, $\|f\|_{L_\infty} \leq C\|f\|_{H^s}$, and $\|\widetilde{\chi}_{\xi}\|_{L_\infty} \leq C\|\widetilde{\chi}_{\xi}\|_{H^s} \leq C\|\widetilde{\chi}_{\xi}\|_{H^m} \leq Ch_{\Xi}^{3-2m}$. Hence, using $\#\Xi \sim q_{\Xi}^{-3} = \rho^3 h_{\Xi}^{-3}$ and taking square roots,

$$\left\|\widetilde{\omega} - \widetilde{\omega}^{\Lambda}\right\|_{\ell_2(\Xi)} \leq Ch_{\Lambda}^{m}h_{\Xi}^{\frac{3}{2}-2m}\|f\|_{H^s}. \tag{6.6.4}$$

For the first term in the parentheses in (6.6.3), note first that

$$\left\|\widetilde{B}^{-1}\widetilde{\omega}\right\|_{\ell_2(\Xi)} \leq \left\|\widetilde{B}^{-1}\right\|_2\|\widetilde{\omega}\|_{\ell_2(\Xi)} \leq Ch_{\Xi}^{-3}\|\widetilde{\omega}\|_{\ell_2(\Xi)}.$$

Using Hölder's inequality, $\#\Xi \sim h_{\Xi}^{-3}$ yet again, and the fact that $\|\widetilde{\chi}_{\xi}\|_{L_2} \leq$

C,

$$\|\widetilde{\omega}\|^2_{\ell_2(\Xi)} = \sum_{\xi\in\Xi} \left| \int_{SO(3)} f \widetilde{\chi}_\xi d\mu \right|^2$$
$$\leq \|f\|^2_{L_2} \sum_{\xi\in\Xi} \|\widetilde{\chi}_\xi\|^2_{L_2}$$
$$\leq C h_\Xi^{-3} \|f\|^2_{L_2}.$$

Therefore, $\left\| \widetilde{B}_\Xi^{-1} \widetilde{\omega} \right\|_{\ell_2(\Xi)} \leq C h_\Xi^{-9/2} \|f\|_{L_2} \leq C h_\Xi^{-9/2} \|f\|_{H^s}$, and so by Corollary 6.5.2

$$\left\| \widetilde{B}_\Xi - \widetilde{B}_\Xi^\Lambda \right\|_2 \left\| \widetilde{B}_\Xi^{-1} \widetilde{\omega} \right\|_{\ell_2(\Xi)} \leq C h_\Xi^{-\frac{11}{2}-2m} h_\Lambda^{m-1} \|f\|_{H^s}. \tag{6.6.5}$$

Putting (6.6.4) and (6.6.5) into (6.6.3), we arrive at

$$\sum_{\eta\in\Xi} \left| \widetilde{\gamma}_\eta - \widetilde{\gamma}_\eta^\Lambda \right|^2 \leq C h_\Xi^{-17-4m} h_\Lambda^{2m-2} \|f\|^2_{H^s}. \tag{6.6.6}$$

For the second factor in (6.6.2), we have by Corollary 6.2.4 and yet again $\#\Xi \sim h_\Xi^{-3}$ that

$$\sum_{\zeta\in\Xi} \|\widetilde{\chi}_\zeta\|^2_{L_2} \leq C h_\Xi^{3-4m} \tag{6.6.7}$$

The result now follows by putting (6.6.6) and (6.6.7) into (6.6.2) and taking square roots. □

Corollary 6.6.8. *Adopt the notation and assumptions from Theorem 6.6.1. There is a constant C such that, for h_Ξ sufficiently small,*

$$\left\| u - \widetilde{u}_\Xi^\Lambda \right\|_{L_2} \leq C \left(h_\Xi^{m-1} + h_\Xi^{-\frac{15}{2}-4m} N^{2-2m} + h_\Xi^{-7-2m} h_\Lambda^{m-1} \right) \|f\|_{H^s}.$$

Remark 6.6.9. We come now to our last main point: we can do as well with quadrature as we can without if h_Λ is small enough. In particular, if

$h_\Lambda \le h_\Xi^p$ with oversampling exponent $p = 3 + \frac{9}{m-1}$, then the quadratization error, $h_\Xi^{-7-2m} h_\Lambda^{m-1}$ will be exactly the Galerkin approximation error, h_Ξ^{m-1} (of course higher oversampling exponents will do even better).

We also must keep in mind the upper bound for h_Λ from Lemma 4.1.25, which is likely more restrictive than the bound given by $h_\Lambda \le h_\Xi^p$. On $SO(3)$ we have $d = 3$, and so the bound is

$$h_\Lambda \le C_0^{\frac{4}{5-2m}} h_\Xi^{1+\frac{16}{2m-5}} 2^{\frac{5}{2}-m}.$$

If indeed the quadrature weights satisfy the lower bound discussed in Remark 4.1.26, then that upper bound is not needed, and the error estimates in Corollary 6.6.8 alone tell us how to choose h_Λ.

Bibliography

[1] N. Aronszajn. La théorie des noyaux reproduisants et ses applications. I. *Proc. Cambridge Philos. Soc.*, 39:133–153, 1943.

[2] N. Aronszajn. Theory of reproducing kernels. *Trans. Amer. Math. Soc.*, 68:337–404, 1950.

[3] T. Aubin. Espaces de Sobolev sur les variétés Riemanniennes. *Bull. Sci. Math. (2)*, 100(2):149–173, 1976.

[4] M. D. Buhmann. Multivariate cardinal interpolation with radial-basis functions. *Constr. Approx.*, 6(3):225–255, 1990.

[5] T. Coulhon, E. Russ, and V. Tardivel-Nachef. Sobolev algebras on Lie groups and Riemannian manifolds. *Amer. J. Math.*, 123(2):283–342, 2001.

[6] M. P. do Carmo. *Differential geometry of curves and surfaces.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1976. Translated from the Portuguese.

[7] M. P. do Carmo. *Riemannian geometry.* Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[8] J. Duchon. Sur l'erreur d'interpolation des fonctions de plusieurs variables par les D^m-splines. *RAIRO Anal. Numér.*, 12(4):325–334, vi, 1978.

[9] G. B. Folland. *Introduction to partial differential equations*. Mathematical Notes. Princeton University Press, Princeton, N.J., 1976. Preliminary informal notes of university courses and seminars in mathematics.

[10] G. B. Folland. *A course in abstract harmonic analysis*. Textbooks in Mathematics. CRC Press, Boca Raton, FL, second edition, 2016.

[11] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright. Localized bases for kernel spaces on the unit sphere. *SIAM J. Numer. Anal.*, 51(5):2538–2562, 2013.

[12] E. Fuselier, T. Hangelbroek, F. J. Narcowich, J. D. Ward, and G. B. Wright. Kernel based quadrature on spheres and other homogeneous spaces. *Numer. Math.*, 127(1):57–92, 2014.

[13] A. Gulisashvili and M. A. Kon. Exact smoothing properties of Schrödinger semigroups. *Amer. J. Math.*, 118(6):1215–1248, 1996.

[14] T. Hangelbroek, F. J. Narcowich, C. Rieger, and J. D. Ward. Direct and inverse results on bounded domains for meshless methods via localized bases on manifolds. In *Contemporary computational mathematics—a celebration of the 80th birthday of Ian Sloan. Vol. 1, 2*, pages 517–543. Springer, Cham, 2018.

[15] T. Hangelbroek, F. J. Narcowich, and J. D. Ward. Kernel approximation on manifolds I: bounding the Lebesgue constant. *SIAM J. Math. Anal.*, 42(4):1732–1760, 2010.

[16] T. Hangelbroek, F. J. Narcowich, and J. D. Ward. Polyharmonic and related kernels on manifolds: interpolation and approximation. *Found. Comput. Math.*, 12(5):625–670, 2012.

[17] T. Hangelbroek and D. Schmid. Surface spline approximation on $SO(3)$. *Appl. Comput. Harmon. Anal.*, 31(2):169–184, 2011.

[18] S. Helgason. *Groups and geometric analysis*, volume 83 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2000. Integral geometry, invariant differential operators, and spherical functions, Corrected reprint of the 1984 original.

[19] A. Hoorfar and M. Hassani. Inequalities on the Lambert W function and hyperpower function. *JIPAM. J. Inequal. Pure Appl. Math.*, 9(2):Article 51, 5, 2008.

[20] R. C. Hoover, A. A. Maciejewski, and R. G. Roberts. Pose detection of 3-d objects using images sampled on so(3), spherical harmonics, and wigner-d matrices. In *2008 IEEE International Conference on Automation Science and Engineering*, pages 47–52, 2008.

[21] W. R. Madych and S. A. Nelson. Polyharmonic cardinal splines. *J. Approx. Theory*, 60(2):141–156, 1990.

[22] F. J. Narcowich, S. T. Rowe, and J. D. Ward. A novel Galerkin method for solving PDEs on the sphere using highly localized kernel bases. *Math. Comp.*, 86(303):197–231, 2017.

[23] F. J. Narcowich, J. D. Ward, and H. Wendland. Sobolev bounds on functions with scattered zeros, with applications to radial basis function surface fitting. *Math. Comp.*, 74(250):743–763, 2005.

[24] F. Riesz and B. Sz.-Nagy. *Functional analysis.* Frederick Ungar Publishing Co., New York, 1955. Translated by Leo F. Boron.

[25] R. Schaback. Native Hilbert spaces for radial basis functions. I. In *New developments in approximation theory (Dortmund, 1998)*, volume 132 of *Internat. Ser. Numer. Math.*, pages 255–282. Birkhäuser, Basel, 1999.

[26] D. Schmid. *Scattered Data Approximation on the Rotation Group and Generalizations.* Berichte aus der Mathematik. Shaker, 2009.

[27] R. S. Strichartz. Analysis of the Laplacian on the complete Riemannian manifold. *J. Functional Analysis*, 52(1):48–79, 1983.

[28] H. Triebel. Spaces of Besov-Hardy-Sobolev type on complete Riemannian manifolds. *Ark. Mat.*, 24(2):299–337, 1986.

[29] H. Wendland. *Scattered data approximation*, volume 17 of *Cambridge Monographs on Applied and Computational Mathematics.* Cambridge University Press, Cambridge, 2005.