Thoughts on Harmonic Analysis on the Sphere: spherical wavelet frames and kernels

vorgelegt von
BMS Phase II, B.Sc.-Mathematiker
Yizhi Sun
ORCID: 0000-0002-2555-5964

an der Fakultät II - Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften
–Dr. rer. nat.–

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. John M. Sullivan
Gutachter: Prof. Dr. Reinhold Schneider
Gutachter: Prof. Dr. Stephan Dahlke

Tag der wissenschaftlichen Aussprache: 27. November 2019

Berlin 2020

Abstract

In the central spirit of harmonic analysis lies the concept of effectively decomposing, analyzing and representing functions or functionals. It has lead to the flourish of Fourier analysis and its modern descendants such as wavelets and its siblings. Especially, the construction of spherical wavelets and its related theory is at junior age.

This dissertation gives a brief summary of existing results in this field, and at the same time create spherical α-wavelets and further develop spherical kernel theory. Among various strategies, two types of spherical wavelets are emphasized, one constructed in the frequency domain, the other generated through stereographic projection. In the former one I discuss localized tight frame design and its directional extension. In the latter one a new anisotropic dilation is defined, and a representation system generated by it consists of the *spherical α-wavelets/shearlets*. Summability properties of those wavelets/shearlets are well established once they are restricted to certain subspaces of square-integrable functions newly defined in this dissertation, including the so called *hollow pole functions*.

Kernels, though deeply rooted in classical theory, can find its variation and application in the frame theory. Indeed, *frame kernel*, a concept which is proposed in this dissertation, is an equivalent formulation to the frame itself. Besides, there exist a variety of kernels which exhibit their own special properties. For zonal kernels, I give its necessary and sufficient conditions to approximate square integrable functions on the sphere. Multiscale kernel, a recently appeared concept, will meet its spherical version here and it turns out to have reproducing property for certain Hilbert space of spherical functions.

One of the climaxes in this work is the invention of two novel frames, based on the two spherical wavelets constructions. In the zonal kernel approach I give frame properties inside the multiresolution structure; while for α-wavelets, I prove that under certain conditions they form tight frames in continuous and discrete setting respectively, following from which are reproducing formulae that enable us to reconstruct or approximate numerically an integrable function or solutions of PDEs. Based on the obtained frames a spherical Galerkin scheme is proposed afterwards. At the end of this dissertation I prove an inner product formula with respect to a recently emerged surface-value dependent inner product space on a triangular mesh, prove its equivalence to the combinatoric inner product, and give eigenvalues estimation for a discrete Laplacian.

Zusammenfassung

Ein zentraler Aspekt der Harmonischen Analysis ist die effiziente Zerlegung, Analyse und Repräsentation von Funktionen und Funktionalen. Diese Konzepte führten zu einem Aufblühen der Fourier Analysis und deren jüngsten Teilgebiete, wie Wavelets und verwandte Methoden. Dabei befindet sich insbesondere die Konstruktion von sphärischen Wavelets und die damit verbundene Theorie noch in den Anfängen ihrer Entwicklung.

Die vorliegende Dissertation gibt einen Überblick über bereits existierende Resultate in diesem Bereich und entwickelt neuartige Instrumente, wie z.B. sphärische Wavelets und sphärische Kernels. Ein besonderes Augenmerk liegt dabei auf zwei Typen von sphärischen Wavelets: Zum einen eine Konstruktion im Frequenzbereich und zum anderen eine Konstruktion durch stereographische Projektionen. Hinsichtlich der Erzeugung im Frequenzbereich werden aktuelle Fortschritte im Bereich des Localized-tight-Frame und deren richtungsabhängige Erweiterungen diskutiert. In den Ausführungen über die Erzeugung von Wavelets durch stereographische Projektionen wird andererseits ein neues Konzept von richtungsabhängigen Dilationen eingeführt, was schließlich zu sogenannten sphärischen α-Wavelets/Shearlets führt. In diesem Zuge können Summierbarkeitseigenschaften hergeleitet werden, nachdem die konstruierten Wavelets/Shearlets auf einen Unterraum von L^2-Funktionen eingeschränkt wurden, was insbesondere sogenannte Hollow-Pole Funktionen einschließt.

Es wird sich herausstellen, dass klassische Kernels ebenfalls sehr vielfältige Anwendungsmöglichkeiten in der modernen Frame-Theorie haben. In diesem Kontext wird das Konzept von Frame-Kernels eingeführt, welches eine äquivalente Formulierung der Frame-Eigenschaft ermöglicht. Darüber hinaus werden noch weitere Beispiele von Kernels hinsichtlich ihrer speziellen Eigenschaften untersucht. Für Zonal-Kernels werden notwendige und hinreichende Bedingungen gezeigt, sodass diese quadratisch integrierbare Funktionen auf der Sphäre approximieren. Schließlich wird eine sphärische Version von Multiskalen-Kernels hergeleitet und es wird gezeigt, dass diese für spezielle Hilberträume von sphärischen Funktionen die Reproduzierbarkeitseigenschaft besitzen.

Ein Hauptergebnis dieser Arbeit bildet die Erfindung von zwei neuartigen Frame-Typen, basierend auf den zwei obengenannten Konstruktionen von sphärischen Wavelets. Hinsichtlich des Zonal-Kernel-Ansatzes wird die Frame-Eigenschaft innerhalb der Multiskalenstuktur nachgewiesen. Für α-Wavelets dahingegen wird bewiesen, dass diese unter bestimmten Annahmen Tight-Frames sowohl im kontinuierlichen als auch im diskreten Sinne bilden. Daraus folgt insbesondere die Reproduzierbarkeitseigenschaft, die die exakte Rekonstruktion von integrierbaren Funktionen oder Lösungen der partiellen Differentialgleichungen ermöglicht. Auf der Grundlage der erhaltenen Frames wird anschließend ein sphärischer Galerkin-Ansatz vorgeschlagen. Zum Abschluss der Arbeit wird eine Formel für das innere Produkt eines kürzlich eingeführten Prä-Hilbertraums bewiesen, der auf einem Dreiecksnetz definiert ist.

Acknowledgements

This dissertation is the fruit of numerous highly efficient working days as well as dreaming days in the library, half seeking inspiration from the classic, half developing into new branches. Even though I never intended to write a big volume, but rather to expose the shining part of an iceberg, both because the full amount of work in this topic is far beyond a doctoral dissertation and exhausting a field that many mathematicians worked for decades is not my purpose, a list of people must be thanked by me personally, so that this independent work will not be solely considered as my work, but rather a natural product out of the support of many. It is dedicated to my family in Hangzhou for their long-lasting care and assistance since childhood. I would give special thanks to Prof. Kutyniok for suggesting me the wonderful dissertation topic about constructing shearlets on the sphere. Some of her suggestions at the beginning of my doctorate, even years later, are still helpful. The writing of Sturm-Liouville theory in the last chapter is inspired during my short visiting at mentor Prof. Fiedler's hospitable group at Freie Universität. My gratitude goes to Prof. Schneider, who takes great responsibility to review my dissertation and guarantees the progress of my doctoral study. I would like to thank Martin and Liselotte for making the German version "Zusammenfassung" possible. My gratitude also belongs to those who appeared to me shortly or frequently like ghosts, no matter to distract me or to encourage me, for they gave me immense inspiration, beauty and love to carry on the research. Contents of this work are either covered and improved selectively by breaking the shell of previous master pieces and extracting their essence, or purely created according to my imagination in order to make a wonderfully integrated hybrid piece; other materials under preparation but not covered in this piece are left to the future. Last but not least I express my gratitude to China Scholarship Council(Grant No.201206320164) and Stipendium des Präsident von Technische Universität Berlin for several years funding support.

Berlin, März 2020

Contents

Chapter 1

Introduction

The purpose of this dissertation is to build analysis tools to analyze information distributed on those surfaces that can be identified with the unit sphere either geometrically or topologically. We shall cover different topics from spherical wavelets, through spherical kernels, to their approximations. However, I am not ambitious in a way to make an encyclopedia here, but rather to explore these areas and their interconnection while at the same time make contribution to some interesting topics along the path, to reflect the beauty and colorfulness lying between harmonic analysis and other mathematical fields and applications.

1.1 Summary of contents

Each chapter is developed systematically and has their own introduction part, hence here I only briefly summarize the contents and contributions. The first chapter serves both as an explanation of mathematical concepts that are crucial in later chapters, such as spherical harmonics, Laplace-Beltrami operator, and as an introduction to spherical operators that emerged recently along with their properties. Especially I derive some commutativity properties of a newly introduced spherical Hilbert-transform. Immediately after that is a reminder of Funk-Hecke formula, integral formulae on the rotation group, as well as the fact that the rotation operator, which plays fundamental role in later chapters, is an irreducible representation of the classical rotation group.

The second chapter starts with the construction of spherical wavelets in the frequency domain, in which a dilation on the half real line using radial basis functions played a key role. In this approach there has been localized tight frames successfully constructed in recent years and an extended version into anisotropic case can be achieved by attaching a steerable directional function onto the radial basis function. In the second section of this chapter, we construct a type of spherical α-wavelet system through stereographic projection, which incorporates spherical wavelets and spherical shearlets. This method is geometrically the most natural, intuitively can be understood as the correspondence between plane and the sphere; it inherits the advantage of conformal mappings, hence has great mathematical properties such as preserving frame structure. What quite different from the previous work of others is that I discuss under a setting incorporating both isotropic and anisotropic dilation. The main new theorem in this chapter gives a pioneering necessary and sufficient condition for a large subclass of square integrable functions to be admissible, based on which further frame properties can be achieved. At the end is a selective review of several other ways of constructing wavelets on the sphere. Each of them has different emphasis on one specific aspect, such as orthogonality, restoration, reduced redundancy, fast computation and so forth.

The third chapter is an extension to different topics with multiple new results that have lines of interconnection beneath. The first topic concerns kernel approximation, in which I would like to invite readers to have a tour through various types of kernels, from the classical reproducing kernel since Hilbert's time to the recently developed multiscale kernels. The underlying questions that I am going to answer in this part include: When is a zonal kernel capable of approximating a square integrable function on the sphere? If such a condition exists, how well can it approximate the function? Does a sphere version multiscale kernel and its corresponding multiscale structure exist? If exists, what properties does it have?

After an affirmative answer to those questions I would like to draw your attention to the construction of spherical frames. In this part two different types of new frame systems are built corresponding to two different ways of spherical wavelet construction in the pervious chapter: one is developed in the frequency domain, where frame property inside the inherited multiresolution structure is for the first time explored, which allows us to transform an arbitrary frame or basis without any good properties into a localized one; while another, technically much more involved, adapts to the wavelets/shearlets systems coming through conformal mapping. Notice that mapping a plane wavelet system to the sphere inevitably suffers from defects such as distortion. In other words, a regular grid on the plane, once we impose on it some translation and dilation operation, does not necessarily give a regular grid on the sphere. Therefore it becomes meaningful to develop frame systems directly on the sphere. In fact, I am going to give both continuous and discrete version localized tight frames on the sphere and their exact reproducing formulae for the stereographic approach, which is absent since twenty years. These new frames, unlike the ones obtained by using energy conservation through conformally projecting planar wavelets or shearlets frames, have the potential to adapt to any preferable grids on the sphere. However, I leave the specific choices of grids, that have been or to be smartly designed and implemented by mathematicians, engineers, scientists and artists, as well as redundancy analysis to future work. Based on those spherical frames and their multiresolution structure a Galerkin scheme which enable us to do numerical analysis of PDEs on the sphere can be formed.

Afterwards a short discussion is devoted to triangulated surface, as another step into the discrete surface world. This part is not tightly related to the previous parts, and by no means I intend to dive deeply here, but rather only give a flavor of how an object living in the continuous world can find their discrete partners. We start by giving a formula for a new inner product of continuous piecewise-linear functions on a given triangulation, comparing it with the weighted inner product on graphs. Along this way, we see two different types of discrete Laplacian on surfaces, one purely combinatoric, while the other is defined in a geometrical manner. The eigenvalues of the geometric discrete Laplacian are less known and I derive their expressions and bounds as a step forward, while the eigenvalues estimation of the combinatoric Laplacian has been well established before this dissertation and I summarize some in the final chapter as supporting material.

The final chapter also contains a self-contained section of elegant introduction to Sturm-Liouville theory is given, from which different kinds of orthogonal polynomials that are essential to the main results in this dissertation are derived naturally. After that, a small section about quaternions and metrics on rotation group gives an alternative representation for the spherical points.

Except for the last chapter which consists of reformulation of known results from personal perspective, most theorems in this dissertation are established and proved either in original or progressive ways, while prepared by discussions based on excellent works of others', so that I believe a balance has been achieved between what are historical and what are innovative in this work.

1.2 Spherical representing systems and operators: old and new

We live in an era that the perception of the world is reformed into mathematical simulation and the understanding of its objects is deepened through systematical deduction using mathematical symbols. Many natural objects like the planet on which we live can be studied abstractly as a sphere, and representing a sphere mathematically is the first step we shall take.

By "representing" I ask two different but united questions. The first question is how to represent points on a sphere? It actually asks, how to build a correspondence relation between the sphere and some parameter space. One obvious way is using the coordinates in Euclidean space, that we have been familiar with in our daily life and in the university analysis and geometry course; another way is to build a group identification, which is partly illustrated in this chapter and the last chapter for reference.

The second question is what are natural representations on the function spaces defined on a sphere? The function space in most context of this work means the L^2 space or subspaces of it, which is Hilbert. We shall encounter several approaches, including representing a function by the superposition of eigenfunctions of the spherical Laplace operator, the integral representation with respect to some wisely chosen kernels, and the representation through expansion of a specially designed representation system.

Let $(I, d\mu)$ be a measure space, a family of elements $\Phi = \{\phi_i\}_{i\in I}$ in a Hilbert space $\mathcal{H}$ is called a dictionary[1] if $\overline{span\{\phi_i\}} = \mathcal{H}$, namely it is a complete subset. Various kinds of dictionaries may be chosen to provide optimal representation of certain class of function spaces, for instance the spherical harmonics introduced below constitute an orthogonal basis for the space of square integrable functions on the sphere, in finite element methods multivariable polynomials under different restrictions provide approximation to Sobolev spaces and wavelets are well adapted to Besov space by its definition. The "optimal" here could either mean an accurate and unique expression of a signal f with fast convergence rate, or it could mean the linear expansion of f in terms of an overcomplete dictionary which has sparse coefficients through minimizing certain norms, or sensitive to special features like high frequency or singularities, depending on objects or tasks; just like preparing a tasty noodle soup for the new years eve, you can either choose ingredients like lamb or fishes inside if the family members

[1] The name dictionary is borrowed from learning theory, with the underlying meaning that the vocabulary inside is sufficiently complete to express any sentences or meanings, in other context it could be alternatively called atoms or molecules, namely a collection of building blocks

are meat fans, or tofu and seegrass if your lovers coincidentally are vegetarian. Being overcomplete intuitively means that to prepare a fish noodle soup, you must at least have fish and noodles, and in addition you can add pepper or soy to achieve different flavors. In real computation usually people aim to obtain a balance between the precision and the sparsity, by minimizing a functional like $\|f - \Phi g\|_2^2 + \alpha M(g)$, with α a regularization parameter and M some cost function.

While there is great flexibility of choosing different ingredients for your soup, you shall not put in everything so that it spoils or does not remains as a fish noodle soup anymore; in other words structure stability and compatibility shall be guaranteed. The concept of frames comes in to provide more flexibility than the orthogonal basis and at the same time requires the basic ingredients that can stably represent a given function. Precisely, a dictionary Φ is called a *frame* if there are positive constants $A \leqslant B$ such that

$$A\|f\|^2 \leqslant \int_I |\langle f, \phi_i \rangle|^2 d\mu \leqslant B\|f\|^2 \tag{1.1}$$

for any $f \in \mathcal{H}$. In particular, when I is discrete, the above inequality becomes

$$A\|f\|^2 \leqslant \sum_{i \in I} |\langle f, \phi_i \rangle|^2 \leqslant B\|f\|^2 \tag{1.2}$$

Meanwhile for any $f \in \mathcal{H}$ there exists at least one dual frame(the canonical dual) $\{\tilde{\phi}_i\}_{i \in I}$ with frame bounds $\frac{1}{B}$ and $\frac{1}{A}$ such that

$$f = \sum_{i \in I} \langle f, \phi_i \rangle \tilde{\phi}_i \tag{1.3}$$

When the frame $\{\phi_i\}_{i \in I}$ is *super tight*, namely when the frame bounds A, B coincide and equal to one, it holds that $\phi_i = \tilde{\phi}_i$ for all $i \in I$. The canonical dual may not be equipped with the properties that Φ has, for instance the dual system does not necessarily have a single generator when the frame Φ does, hence does not inherits a wavelet structure. In fact, there might exist infinitely many (alternate) duals, but how to choose a dual wisely is much depending on the problems to deal with.

Spherical harmonics form an orthogonal system, hence its dual is itself. It has become a useful tool to analyze functions on the sphere, since the time of Laplace and Legendre. A spherical harmonic Y_l is a homogeneous polynomial of degree l which solves the Laplace equation. After being restricted on the sphere, they are sometimes called surface spherical harmonics. In this work, if we don't give additional clarification, it is assumed that spherical harmonics are already restricted and they satisfy the equations

$$\Delta_{\mathbb{S}^2} Y_l = -l(l+1)Y_l \tag{1.4}$$

where

$$\Delta_{\mathbb{S}^2} = \frac{1}{\sin\theta}\Big(\frac{\partial}{\partial\theta}(\sin\theta\frac{\partial}{\partial\theta}) + \frac{1}{\sin\theta}\frac{\partial^2}{\partial\varphi^2}\Big) \tag{1.5}$$

is the Laplace-Beltrami Operator on $\mathbb{S}^2$. The *operator of infinitesimal rotations*[65] or *operators of angular momentum*[71] up to a change of sign i are

$$\begin{aligned}
L_x &= -i(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}) = i(\sin\varphi\frac{\partial}{\partial\theta} + \cot\theta\cos\varphi\frac{\partial}{\partial\varphi}) \\
L_y &= -i(z\frac{\partial}{\partial x} - x\frac{\partial}{\partial z}) = i(-\cos\varphi\frac{\partial}{\partial\theta} + \cot\theta\sin\varphi\frac{\partial}{\partial\varphi}) \\
L_z &= -i(x\frac{\partial}{\partial y} - y\frac{\partial}{\partial x}) = -i\frac{\partial}{\partial\varphi}
\end{aligned} \tag{1.6}$$

It can be immediately verified that

$$-\Delta_{\mathbb{S}^2} = L_x^2 + L_y^2 + L_z^2 \tag{1.7}$$

If we denote by $(\eta_1, \eta_2, \eta_3)^T$ the eigenvector of $\sigma \in SO(3)$ with its length equal to the rotation angle ϕ and the identity matrix of rotation group by σ_0, when ϕ is mall, there is

$$\mathcal{R}(\sigma) = \mathcal{R}(\eta_1, \eta_2, \eta_3) = \mathcal{R}(\sigma_0) - iL_x\eta_1 - iL_y\eta_2 - iL_z\eta_3 + O(\phi^2)$$

where $\mathcal{R}(\sigma_0)$ is just the identity operator.

Since for $s, s^{'} > 0$ there is the obvious relation

$$\mathcal{R}(s\eta_1, s\eta_2, s\eta_3)\mathcal{R}(s^{'}\eta_1, s^{'}\eta_2, s^{'}\eta_3) = \mathcal{R}((s+s^{'})\eta_1, (s+s^{'})\eta_2, (s+s^{'})\eta_3)$$

it follows that

$$\frac{d\mathcal{R}(s\eta_1, s\eta_2, s\eta_3)}{ds} = -i\mathcal{R}(s\eta_1, s\eta_2, s\eta_3)L_{\vec{\eta}} \tag{1.8}$$

where $L_{\vec{\eta}} = L_x\eta_1 + L_y\eta_2 + L_z\eta_3$, hence

$$e^{-iL_{\vec{\eta}}} = \mathcal{R}(\eta_1, \eta_2, \eta_3) \tag{1.9}$$

and the unitarity of representation $\mathcal{R}(\eta_1, \eta_2, \eta_3)$ is equivalent to the hermiticity of the operators L_x, L_y and L_z in $L^2(\mathbb{S}^2)$.

Notice that the element $\tilde{\sigma} = \sigma_1 \sigma \sigma_1^{-1}$ leaves $\sigma_1(\eta_1, \eta_2, \eta_3)^T$ invariant. Furthermore, if $\eta^\perp$ is a unit vector perpendicular to $\eta = (\eta_1, \eta_2, \eta_3)^T$ such that $(\sigma\eta^\perp, \eta^\perp) = \cos\phi$, then $(\tilde{\sigma}\sigma_1\eta^\perp, \sigma_1\eta^\perp) = \cos\phi$. Thus the rotation angle of $\tilde{\sigma}$ is the same as σ. Assume that σ_1 is a small rotation, namely

$$\mathcal{R}(\sigma_1) = \mathcal{R}(\sigma_0) - i(L_x\zeta_1 + L_y\zeta_2 + L_z\zeta_3) + O(|\zeta|^2)$$

with $|\zeta|$ small. Let $\eta = (|\eta|, 0, 0)^T$ and $\zeta = (0, |\zeta|, 0)^T$, then as a consequence of the identity $\mathcal{R}(\tilde{\sigma}) = \mathcal{R}(\sigma_1)\mathcal{R}(\sigma)\mathcal{R}(\sigma_1^{-1})$ there is

$$\begin{aligned}\mathcal{R}(\tilde{\sigma}) &= \mathcal{R}(\sigma_0) - i(L_x\tilde{\eta}_1 + L_y\tilde{\eta}_2 + L_z\tilde{\eta}_3) + O(|\eta|^2) \\ &= (\mathcal{R}(\sigma_0) - i|\zeta|L_y)(\mathcal{R}(\sigma_0) - i|\eta|L_x)(\mathcal{R}(\sigma_0) - i|\zeta|L_y) \\ &= \mathcal{R}(\sigma_0) - i|\eta|L_x + |\eta| \cdot [L_x, L_y]|\zeta| + O(|\zeta|^2)\end{aligned} \tag{1.10}$$

Identity $\sigma_1\zeta = \zeta$ implies

$$\sigma_1 = \begin{pmatrix} \cos\phi_1 & 0 & \sin\phi_1 \\ 0 & 1 & 0 \\ -\sin\phi_1 & 0 & \cos\phi_1 \end{pmatrix}$$

where $\phi_1 = |\zeta|$ and $\tilde{\eta} = \sigma_1\eta = (|\eta|\cos\phi_1, 0, -|\eta|\sin\phi_1)^T$, hence by comparing the first order term of $|\zeta|$ in (1.10) we obtain $[L_x, L_y] = iL_z$. By exchanging the role of x, y and z, we arrive at the following elegant commutation rules

$$[L_x, L_y] = iL_z, \quad [L_y, L_z] = iL_x, \quad [L_z, L_x] = iL_y \tag{1.11}$$

By defining

$$L_+ = L_x + iL_y, \quad L_- = L_x - iL_y$$

relations (1.11) become

$$[L_+, L_z] = -L_+, \quad [L_-, L_z] = L_-, \quad [L_+, L_-] = 2L_z \tag{1.12}$$

Notice that L_+, L_- and L_z are self-adjoint and if $m \in \mathbb{R}$ nonzero is an eigenvalue of L_z with eigenfunction f, then L_+f is an eigenvector corresponding to eigenvalue $m+1$ while L_-f is one corresponding to eigenvalue $m-1$. Since the total number of distinct eigenvalues is finite, assume that the largest of them is l and write its normalized eigenfunction as Y^l, similarly the smallest of them is written as l^* with eigenfunction Y^{l^*}.

Define by induction $\underline{\alpha_m}Y^{m-1} = L_-Y^m$ where $\underline{\alpha_m}$ are chosen so that

$$\langle Y^{m-1}, Y^{m-1}\rangle = 1$$

Since $L_+Y^l = 0$, that is

$$L_+Y^{l-1} = \frac{1}{\underline{\alpha_l}} L_+L_-Y^l = \frac{2}{\underline{\alpha_l}} L_zY^l = \frac{2l}{\underline{\alpha_l}} Y^l = \overline{\alpha_l} Y^l \tag{1.13}$$

an induction argument leads to the claim that L_+Y^m is proportional to Y^{m+1}, namely $L_+Y^m = \overline{\alpha_{m+1}} Y^{m+1}$. Moreover, the fact that $L_+^\dagger = L_-$ gives

$$\overline{\alpha_m} \langle Y^m, Y^m \rangle = \underline{\alpha_m} \langle Y^{m-1}, Y^{m-1} \rangle$$

namely $\overline{\alpha_m} = \underline{\alpha_m} = \alpha_m$. Thus from the observation that

$$\alpha_{m+1}^2 Y^{m+1} = L_+L_-Y^{m+1} = (2L_z + L_-L_+)Y^{m+1} = (2m + 2 + \alpha_{m+2}^2)Y^{m+1}$$

we obtain the relation $\alpha_{m+1}^2 - \alpha_m^2 = -2m$. From this relation, our observation that $\alpha_l^2 = 2l$ leads to the expression

$$\alpha_m^2 = (l + m)(l - m + 1) \tag{1.14}$$

$L_-Y^{l^*} = 0$ gives $\alpha_{l^*} = 0$, hence $l^* = -l$. Therefore l is either an integer or half an odd number and every irreducible representation is uniquely determined by l with its dimension equal to $2l + 1$.

In real calculation and approximation, frequently used is the normalized expression

$$Y_l^m(\theta, \varphi) = (-1)^m \sqrt{\frac{(2l + 1)(l - m)!}{4\pi(l + m)!}} e^{im\varphi} P_l^m(\cos\theta) \tag{1.15}$$

which naturally appears when one solves the Laplace equation through separation of variables, where $l \in \mathbb{N}$ and $|m| \leqslant l$. Those spherical harmonics of degree l form an orthonormal basis for the space of homogeneous polynomials of degree l, denoted by $\mathcal{H}_l$.

From the property that $-i\frac{\partial}{\partial\varphi} Y_l^m = mY_l^m$ we already see that $Y_l^m(\varphi, \theta) = e^{im\varphi} F_l^m(\theta)$ for some function $F_l^m(\theta) =: \tilde{F}_l^m(\cos\theta)$. Let $x = \cos\theta$, from (1.5) we deduce that

$$(1 - x^2)\frac{d^2}{dx^2}\tilde{F}_l^m(x) - 2x\frac{d}{dx}\tilde{F}_l^m(x) + \left[l(l + 1) - \frac{m^2}{1 - x^2}\right]\tilde{F}_l^m(x) = 0 \tag{1.16}$$

when $m = 0$ it is exactly differential equation (4.41), thus $\tilde{F}_l^m$ are nothing else but the associated Legendre Polynomials

$$\begin{aligned}
P_l^m(x) &:= (-1)^m (1-x^2)^{m/2} \frac{d^m}{dx^m} P_l(x) \\
&= \frac{(-1)^m (1-x^2)^{m/2}}{2^l l!} \frac{d^{l+m}}{dx^{l+m}} (x^2-1)^l \\
&= (-1)^m \frac{(l+m)!}{(l-m)!} P_l^{-m}(x)
\end{aligned} \tag{1.17}$$

which satisfy the orthogonality relation that

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x) dx = \frac{(l+m)!}{(l-m)!} \frac{2}{2l+1} \delta_{l,l'} \tag{1.18}$$

where the last equality of (1.17) follows from Rodrigues formula

$$\frac{d^{l-m}(x^2-1)^l}{dx^{l-m}} = \frac{(l-m)!}{(l+m)!} (x^2-1)^m \frac{d^{l+m}(x^2-1)^l}{dx^{l+m}}$$

Sometimes Laplace's integral expression for Legendre polynomials is as useful as the differential expression. It reads equivalently as

$$P_l(x) = \frac{1}{\pi} \int_0^{\pi} \left(x \pm \sqrt{x^2-1} \cos\theta\right)^l d\theta \tag{1.19a}$$

$$P_l(x) = \frac{1}{\pi} \int_0^{\pi} \left(x \mp \sqrt{x^2-1} \cos\theta\right)^{-l-1} d\theta \tag{1.19b}$$

More generally there is

$$P_l^m(x) = \gamma_{l,m}^{+} \int_0^{\pi} \left(x \mp \sqrt{x^2-1} \cos\theta\right)^l \cos(m\theta) d\theta$$

$$P_l^m(x) = \gamma_{l,m}^{-} \int_0^{\pi} \left(x \pm \sqrt{x^2-1} \cos\theta\right)^{-l-1} \cos(m\theta) d\theta$$

where $\gamma_{l,m}^{+} = (\pm 1)^m \frac{(l+m)!}{\pi l!} e^{-\frac{m}{2}\pi i}$ and $\gamma_{l,m}^{-} = (\pm 1)^m \frac{l!}{\pi(l-m)!} e^{-\frac{m}{2}\pi i}$.

The spherical Hilbert transform, as an analogue of the plane situation, firstly appears in [85](but no mathematical properties are given there)

$$\widehat{(\mathfrak{H} f)}_{lm} = \begin{cases} -i\widehat{f}_{lm} & m > 0 \\ 0 & m = 0 \\ i\widehat{f}_{lm} & m < 0 \end{cases} \tag{1.21}$$

where $\widehat{f_{lm}}$ is the Fourier coefficient of a square integrable function f with respect to the spherical harmonics Y_l^m. Let us derive some new commutativity relations and formulae of those operators in the next theorem.

Theorem 1.1. *(i) $i\mathfrak{H}$ is self-adjoint and $(i\mathfrak{H})^2 f = -f$ for any $f \in L^2(\mathbb{S}^2)$. When $\widehat{f}_{l,m} = \widehat{f}_{l,-m}$ for all l and $|m| \leqslant l$, f and $i\mathfrak{H}f$ are orthogonal.*

(ii) The differential operator L_x, L_y, L_z defined in (1.6) commute with $\mathfrak{H}$. In particular $[\Delta_{\mathbb{S}^2}, \mathfrak{H}] \equiv 0$.

(iii) Given any bounded operator $A : L^2(\mathbb{S}^2) \to L^2(\mathbb{S}^2)$, there is

$$[\Delta_{\mathbb{S}^2}, A]\mathcal{H}_l \perp \mathcal{H}_l \quad \textit{and} \quad [\Delta_{\mathbb{S}^2}, A]\mathfrak{M}_m \perp \mathfrak{M}_m \tag{1.22}$$

where for each $m \in \mathbb{Z}$,

$$\mathfrak{M}_m = \overline{span\{Y_l^m : l \geqslant 0\}} \tag{1.23}$$

(iv) For each l, there holds the identity

$$\begin{aligned} &\{\widehat{(\mathfrak{H}f)}_{l,m}\}_m \circledast \{\widehat{(\mathfrak{H}f)}_{l,m}\}_m(n) - \{\widehat{f}_{l,m}\}_m \circledast \{\widehat{f}_{l,m}\}_m(n) \\ &= -isgn([n]) \left(\{\widehat{(\mathfrak{H}f)}_{l,m}\}_m \circledast \{\widehat{f}_{l,m}\}_m(n) + \{\widehat{f}_{l,m}\}_m \circledast \{\widehat{(\mathfrak{H}f)}_{l,m}\}_m(n) \right) \end{aligned} \tag{1.24}$$

where $(\{a_m\} \circledast \{b_m\})(n) = \sum_{|m| \leqslant l} a_{[n-m]} b_m$ and $[n]$ taking the module[2].

Proof. The claim that $i\mathfrak{H}$ is self-adjoint and $(i\mathfrak{H})^2 f = -f$ is obvious. The rest of (i) is a result of the observation that

$$\langle f, i\mathfrak{H}f \rangle_{L^2} = \sum_l \sum_{|m| \leq l} sgn(m) |\hat{f}_{lm}|^2 \tag{1.25}$$

is zero since each m term cancels the $-m$ term.

It can be checked through calculation that the identity

$$\begin{aligned} L_z \mathfrak{H} Y_l^m &= \begin{cases} -imY_l^m & \text{if } m > 0 \\ 0 & \text{if } m = 0 \\ imY_l^m & \text{if } m < 0 \end{cases} \\ &= -i|m|Y_l^m \\ &= \mathfrak{H} L_z Y_l^m \end{aligned}$$

[2]Here I make the convention that $[n] = -(|n| \mod l+1)$ when n is a negative integer; $[n] = l+1$ when $(n - m \mod l+1) = 0$

holds for any spherical harmonics, hence for square integrable functions. Observe that

$$\begin{aligned}&\mathfrak{H}L_+Y_l^m = -i\overline{\alpha_{m+1}}\mathrm{sgn}(m)Y_l^m = L_+\mathfrak{H}Y_l^m \quad \text{for} \quad m \neq l\\ &\mathfrak{H}L_+Y_l^l = 0 = L_+\mathfrak{H}Y_l^l\end{aligned}$$

and $[\mathfrak{H}, L_-]Y_l^m = 0$ for any m. As linear combination of L_+ and L_-, operators L_x and L_y commute with $\mathfrak{H}$ respectively. The assertion that $\Delta_{\mathbb{S}^2}$ commutes with $\mathfrak{H}$ is thus an immediate result of (1.7).

Besides, there is

$$\begin{aligned}\langle[\Delta_{\mathbb{S}^2}, A]Y_l^m, Y_{l'}^{m'}\rangle &= \langle AY_l^m, \Delta_{\mathbb{S}^2}Y_{l'}^{m'}\rangle - \langle A\Delta_{\mathbb{S}^2}Y_l^m, Y_{l'}^{m'}\rangle\\ &= [-l'(l'+1) + l(l+1)]\langle AY_l^m, Y_{l'}^{m'}\rangle\end{aligned} \tag{1.26}$$

Similarly,

$$\begin{aligned}\langle[L_z, A]Y_l^m, Y_{l'}^{m'}\rangle &= -\langle AY_l^m, L_zY_{l'}^{m'}\rangle - \langle AL_zY_l^m, Y_{l'}^{m'}\rangle\\ &= (m' - m)\langle AY_l^m, Y_{l'}^{m'}\rangle\end{aligned} \tag{1.27}$$

hence (iii) is verified.

Finally, the left hand side of (1.24) is equal to

$$-\sum_{|m|\leqslant l} (\mathrm{sgn}([n-m])\mathrm{sgn}(m) + 1)\, \hat{f}_{l,[n-m]}\hat{f}_{l,m}$$

while the right hand side is equal to

$$-\mathrm{sgn}([\mathrm{n}]) \sum_{|m|\leqslant l} (\mathrm{sgn}([n-m]) + \mathrm{sgn}(m))\, \hat{f}_{l,[n-m]}\hat{f}_{l,m}$$

hence identical to each other. Indeed, under our convention if $[n-m]$ and m are of opposite sign, both left and right hand side vanish; if $\mathrm{sgn}([n-m]) = \mathrm{sgn}(m) = 1$, then $\mathrm{sgn}[n] = 1$; if $\mathrm{sgn}([n-m]) = \mathrm{sgn}(m) = -1$, then $\mathrm{sgn}[n] = -1$. □

The renormalized Poisson kernel has the expression

$$Q_r(t) = \frac{1}{4\pi}\frac{1-r^2}{(1-2rt+r^2)^{3/2}} = \sum_{l=0}^{\infty}\frac{2l+1}{4\pi}r^lP_l(t) \tag{1.28}$$

with $r \in (0,1)$ and $t \in [-1,1]$. It has the following approximation property for continuous functions, which shall be used later. Its proof can be found in [40] for instance, but for completeness we give a greatly simplified version.

Lemma 1.2. *For any continuous function f on the two sphere, there is*

$$\lim_{r\to 1^-}\sup_{y\in\mathbb{S}^2}\left|\int_{\mathbb{S}^2} Q_r(x\cdot y)f(x)d\Omega(x)-f(y)\right|=0$$

Proof. Firstly notice that for any $\epsilon>0$ there exists $\delta\in(0,1)$ such that $|f(x)-f(y)|<\epsilon$ whenever $|x\cdot y-1|<\delta$.

$$\int_{\mathbb{S}^2} Q_r(x\cdot y)dy=\sum_{n=0}^{\infty}\frac{2n+1}{2}r^n\int_{-1}^{1}P_n(t)dt=1$$

hence

$$\begin{aligned}
&\left|\int_{\mathbb{S}^2} Q_r(x\cdot y)f(y)dy-f(x)\right|\\
&\leqslant\int_{\mathbb{S}^2} Q_r(x\cdot y)|f(y)-f(x)|dy\\
&\leqslant 2\|f\|_C\int_{x\cdot y\leqslant 1-\delta} Q_r(x\cdot y)dy+\frac{\epsilon}{2}\int_{x\cdot y>1-\delta} Q_r(x\cdot y)dy\\
&\leqslant \|f\|_C\int_{-1}^{1-\delta}\frac{1-r^2}{(1+r^2-2rt)^{3/2}}dt+\frac{\epsilon}{2}\int_{1-\delta}^{1}\frac{1-r^2}{(1+r^2-2rt)^{3/2}}dt\\
&=\frac{1-r^2}{r}\left.\frac{\|f\|_C}{\sqrt{1+r^2-2rt}}\right|_{-1}^{1-\delta}+\frac{1-r^2}{r}\left.\frac{\epsilon}{2\sqrt{1+r^2-2rt}}\right|_{1-\delta}^{1}\\
&\to\frac{\epsilon}{2}
\end{aligned}$$

as r approaches 1^-. Due to the arbitrariness of ϵ, the claim follows.

$\square$

Spherical harmonics in (1.15) have the important property that they form an orthonormal basis for the Hilbert space $L^2(\mathbb{S}^2,d\Omega)$. Therefore

$$Z_l(\xi,\eta)=\sum_{|m|\leq l}\overline{Y_l^m(\xi)}Y_l^m(\eta) \tag{1.29}$$

gives a unique reproducing kernel of the space $\mathcal{H}_l$. This definition does not depend on the orthonormal system that we choose. In fact, the well known addition theorem says that

$$Z_l(\xi,\eta)=\frac{2l+1}{4\pi}P_l(\xi\cdot\eta). \tag{1.30}$$

where P_l are Legendre Polynomials. For curious readers, I refer to [10] or [36] for its proof. In a general dimension, $\mathbb{P}_l^d$ can be defined in the same way, with

$$Z_l^d = \frac{l+\lambda_d}{|\mathbb{S}^{d-1}|\lambda_d} P_l^{\lambda_d} \tag{1.31}$$

where $\lambda_d = (d-2)/2$ and $P_l^{\lambda_d}$ is the Gegenbauer polynomials. Define projection operators $\mathbb{P}_l : L^2(\mathbb{S}^2) \to \mathcal{H}_l$ as

$$\mathbb{P}_l^d f(\xi) := f * Z_l^d(\xi) := \int_{\mathbb{S}^{d-1}} Z_l^d(\xi \cdot \eta) f(\eta) d\Omega(\eta) \tag{1.32}$$

In most of the discussion below we assume $d = 3$ and briefly denote it by $\mathbb{P}_l$.

Every function on $[-1, 1]$ which is integrable with respect to the weight function $(1-t^2)^{\frac{d-3}{2}}$ satisfies the so called Funk-Hecke formula,

$$\int_{\mathbb{S}^2} f(\xi \cdot \eta) Y(\eta) d\Omega(\eta) = c_l Y(\xi) \tag{1.33}$$

with $c_l = |\mathbb{S}^{d-1}| \int_{-1}^{1} f(t)(1-t^2)^{\frac{d-3}{2}} \frac{P_l^{\lambda_d}(t)}{P_l^{\lambda_d}(1)} dt$ and $Y \in \mathcal{H}_l$. Furthermore, as an immediate consequence,

$$\int_{\mathbb{S}^2} Z_l(\xi, \omega) Z_l(\eta, \omega) d\Omega(\omega) = Z_l(\xi, \eta) \tag{1.34}$$

Let $SO(3)$ be the rotation group, consisting of matrices that are orthogonal and of determinant one. This group or its representation plays an essential role in the analysis on spheres. For instance, the Fourier transform of $f \in L^1(\mathbb{S}^2)$ can be defined alternatively as

$$\mathcal{F} f(\gamma) = \int_{SO(3)/SO(2)} f(x)(-x, \gamma) d\mu(x), \ \gamma \in \Gamma \tag{1.35}$$

where Γ is the dual group of $SO(3)/SO(2)$, annihilator of $SO(2)$.

Since $SO(3)/SO(2)$ is compact, Γ is discrete. Let $q : SO(3) \to SO(3)/SO(2)$ be the natural homomorphism, then

$$f \to \int_{SO(3)} f \circ q \ d\sigma \tag{1.36}$$

is a bounded linear functional on $C(SO(3)/SO(2))$, hence induces a unique measure $\mu \in M(SO(3)/SO(2))$ with $\|\mu\| \leq \|\sigma\|$ and

$$\int_{SO(3)} f \circ q \ d\sigma = \int_{SO(3)/SO(2)} f d\mu \tag{1.37}$$

The operator defined by $Sf([x]) = \int_{SO(2)} f(xy)d\mu(y)$ projects functions in $C(SO(3))$ onto $C(SO(3)/SO(2))$.

Besides, the integral on $SO(3)$ can be expressed explicitly in the form

$$\int_{SO(3)} f d\sigma = \frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} f(\bar{\varphi}_1,\bar{\theta},\bar{\varphi}_2)\sin\bar{\theta} d\bar{\varphi}_1 d\bar{\theta} d\bar{\varphi}_2 \tag{1.38}$$

where $(\bar{\varphi}_1,\bar{\theta},\bar{\varphi}_2)$ is the Euler angle. When f depends only on the rotation angle α it has the expression

$$\int_{SO(3)} f d\sigma = \frac{2}{\pi}\int_0^{\pi} f(\alpha)\sin^2\frac{\alpha}{2}d\alpha \tag{1.39}$$

where $\alpha(g_1 g_2^T) = \arccos(\frac{1}{2}(Trg_1 g_2^T - 1))$ is the rotation angle.

What is important to later chapters is the left regular unitary representation of $SO(3)$ on the Hilbert space $L^2(\mathbb{S}^2)$ defined by

$$(\mathcal{R}(\sigma)f)(\omega) = f(\sigma^{-1}\omega) \tag{1.40}$$

where $\sigma \in SO(3)$. $\mathcal{R}(\sigma)$ is called *rotation operator*. [71] is a brilliant reference for some of its interesting properties and applications in angular momentum theory of quantum mechanics. [16] is a classic in both group representation theory and orthogonal polynomials. The rotation operator, together with the spherical dilation operators that are going to be defined and discussed intensively in the next chapter, are building blocks for the frame systems in this dissertation.

We have mentioned that for nonnegative integer l, $\{Y_l^{-l},\cdots,Y_l^l\}$ form an orthonormal basis for $\mathcal{H}_l$. In this situation $\mathcal{R}$ is irreducible, for otherwise suppose there is an invariant subspace $\mathcal{H}'$, then it is invariant under L_+, L_- and L_z as well. Take any $h' = \sum_{m=-l}^{l} c_m Y_l^m \in \mathcal{H}'$, let m' be the smallest index such that $c_{m'} \neq 0$. There is $L_+^{l-m'}h' = c_{m'}\alpha_l\cdots\alpha_{m'+1}Y_l^l \in \mathcal{H}'$ by invariance under L_+. Thus $Y_l^l \in \mathcal{H}'$ and $L_-^{l-m}Y_l^l = \alpha_l\alpha_{l-1}\cdots\alpha_{m+1}Y_l^m \in \mathcal{H}'$ implies that $Y_l^m \in \mathcal{H}'$ for arbitrary $|m| \leqslant l$, hence $\mathcal{H}' = \mathcal{H}_l$.

Similarly we can define the regular representation of $SO(d)$ and denote by the $\mathcal{R}^{l,d}$ its restriction on $H_{d,l}$. In fact, the next classic theorem generalize the irreducible property to $\mathcal{R}^{d,l}$ on $SO(d)$. Its proof is quite instructive in decomposing polynomial spaces aspect, hence I include it here. However, it is like a cherry on the cake, not every child likes or needs to eat it.

Theorem 1.3. *$\mathcal{R}^{d,l}$ is an irreducible representation of $SO(d)$.*

Proof. For $\sigma_{d-1} \in SO(d-1)$, there is

$$\mathcal{R}^{l,d}(\sigma_{d-1}) = \bigoplus_{l' \leqslant l} \mathcal{R}^{l',d-1}(\sigma_{d-1}) \tag{1.41}$$

Indeed, for $f := x_d^{l-l'} h^{l'} \in x_d^{l-l'} \mathcal{H}_{d-1,l'}$ assume with out loss of generality that σ_{d-1} leaves e_d unchanged, then there is

$$\mathcal{R}^{l,d}(\sigma_{d-1}) f = x_d^{l-l'} h^{l'} (\sigma_{d-1}^{-1}(x_1, \cdots, x_{d-1}))$$

hence $x_d^{l-l'} \mathcal{H}_{d-1,l'}$ is an invariant subspace of $\mathcal{P}_{d,l}$, the space of homogeneous polynomials of degree l. Since $\mathcal{P}_{d,l} = \mathcal{H}_{d,l} \oplus r^2 \mathcal{P}_{d,l-2}$, there is

$$\mathcal{P}_{d,l} = \bigoplus_{k=0}^{\llcorner \frac{l}{2} \lrcorner} r^{2k} \mathcal{H}_{d,l-2k} \tag{1.42}$$

Let $q : \mathcal{P}_{d,l} \mapsto \mathcal{P}_{d,l}/r^2\mathcal{P}_{d,l-2}$ be the canonical map and $V_{l,l'}^d = q(x_d^{l-l'} \mathcal{H}_{d-1,l'})$ which is invariant under $\mathcal{R}^{l,d}(\sigma_{d-1})$. By (1.42) there is

$$\begin{aligned}
\mathcal{P}_{d,l} &= r^2 \mathcal{P}_{d,l-2} + x_d \, \mathcal{P}_{d-1,l-1}|_{x'} + \mathcal{P}_{d-1,l}|_{x'} \\
&= r^2 \mathcal{P}_{d,l-2} + x_d \sum_{k=0}^{\llcorner \frac{l-1}{2} \lrcorner} (r^2 - x_d^2)^k \mathcal{H}_{d-1,l-1-2k} + \sum_{k=0}^{\llcorner \frac{l}{2} \lrcorner} (r^2 - x_d^2)^k \mathcal{H}_{d-1,l-2k} \\
&\subset r^2 \mathcal{P}_{d,l-2} + \bigoplus_{k=0}^{l} x_d^k \mathcal{H}_{d-1,l-k}
\end{aligned}$$

where $x' = (x_1, \cdots, x_{d-1})$, and the converse inclusion is obvious. Thus we get

$$\mathcal{P}_{d,l} = r^2 \mathcal{P}_{d,l-2} + \bigoplus_{k=0}^{l} x_d^k \mathcal{H}_{d-1,l-k} \tag{1.43}$$

hence $q(\mathcal{P}_{d,l}) = \mathcal{P}_{d,l}/r^2\mathcal{P}_{d,l-2} = \bigoplus_{k=0}^{l} V_{l,k}^d$ and (1.41) follows.

As a consequence is the irreducibility of $\mathcal{R}^{l,d}$ of the group $SO(d)$ for $d \geqslant 3$. Suppose this has been proved for $d \leqslant n-1$. For $d = n$, notice in (1.41) that the restriction of $\mathcal{R}^{l,d}$ on $SO(d-1)$ is the the direct sum of irreducible $\mathcal{R}^{l',d-1}$ on $V_{l,l'}^d$. Therefore if W is a non-trivial invariant subspace under $\mathcal{R}^{l,d}$, it must be of the form $W = \bigoplus_{l' \in \mathcal{L}} V^d_{l,l'}$ for some subset $\mathcal{L}$ of $\{0, \cdots, l\}$. Thus it only remains to prove $\mathcal{L} = \{0, \cdots, l\}$.

Let l_s be the smallest index in $\mathcal{L}$ and l_b biggest and assume $V^d_{l,l'} \subset W$. Since each rotation can be decomposed as a series of rotations in planes x_j, x_k $(j \neq k)$, infinitesimal operators $L_{j,k} = x_k \frac{\partial f}{x_j} - x_j \frac{\partial f}{\partial x_k}$ must leave $V^d_{l,l'}$ invariant and choose some j such that $\frac{\partial h^{d-1,l_b}}{\partial x_j} \neq 0$. If neither j nor k are equal to d, then $L_{jk}(x_d^{l-l'} h^{d-1,l'}(x')) \in x_d^{l-l'} \mathcal{H}_{d-1,l'}$. If $k = d$ for instance(the same to the situation $j = d$),

$$
\begin{aligned}
L_{j,d}(x_d^{l-l'} h^{d-1,l'}(x')) &= x_d^{l-l'+1} \frac{\partial h^{d-1,l'}}{\partial x_j} - (l-l')x_d^{l-l'-1} x_j h^{d-1,l'} \\
&\in x_d^{l-l'+1} \mathcal{H}_{d-1,l'-1} + x_d^{l-l'-1} \mathcal{H}_{d-1,l'+1}
\end{aligned}
$$

Thus if $l_b < l$, then $L_{j,d}(x_d^{l-l_b} \mathcal{H}_{d-1,l_b}) \cap x_d^{l-l_b-1} \mathcal{H}_{d-1,l_b+1} \neq \emptyset$ and

$$x_d^{l-l_b-1} \mathcal{H}_{d-1,l_b+1} \subset W$$

since W is invariant, contradictory to the assumption that l_b is the biggest index. Similarly, if $l_s > 0$, then $\frac{\partial h^{d-1,l_s}}{\partial x_j} \neq 0$ implies $x_d^{l-l_s+1} \mathcal{H}_{d-1,l_s-1} \subset W$, contradictory to the smallest assumption on l_s. Therefore we can conclude that $W = \mathcal{H}_{d,l}$.

$\square$

On the one hand we have been immersed in the exhilarating success of spherical harmonics which not only expand the L^2 space, but also form rotational invariant subspaces with hierarchical structure; on the other hand we have to admit that, the nowadays most widely used planar wavelets, which was originally based on the idea of Gabor functions, have replaced traditional Fourier transform in dealing with shock waves in seismology, acoustic or image signals characterized with singularities, and spherical harmonics face the same awkwardness in this aspect due to its global feature. Furthermore, after witnessing the fast development of wavelets' planar descendants like ridgelets, curvelets, brushlets, contourlets and cone-adapted shearlets in an attempt to complement the inefficiency of the wavelets in detecting anisotropic structures, it is very natural to ask for the generalization of those "-lets" adapting to an arbitrary surface.

Chapter 2

Spherical Dilation Systems

In the past decade emerged multiple methods to define wavelets on the sphere or a general manifold, and each of them bears different merits. Furthermore, like those affine-like systems on the plane, whenever possible, it is convenient to introduce operations like translation and dilation. Not only because through them a system can be generated and implemented in a simple manner, but also pertaining to them there are wonderful properties. Thus in this section let us pay attention to those special wavelet systems equipped with various means of dilation operation.

2.1 Dilation in frequency domain

One attempt of this, the so called the curvelets on the sphere, is given in [22]. The idea behind this is similar to that of differential geometry, namely the surface is divided smoothly into charts for multiple scales and then apply the curvelets on each of the building block. The scaling function there is proposed as follows. Let $L = 2^J$ for some $J \in \mathbb{N}$ and

$$\phi_L = \sum_{l=0}^{L} \widehat{\phi}_L(l,0) Y_{l,0} \tag{2.1}$$

to form a sequence of functions of multi-scale ϕ_L, $\phi_{2^{-1}L}$, $\cdots$, $\phi_{2^{-j}L}$.

With the low pass filter

$$\widehat{h}_j(l,m) = \begin{cases} \frac{\widehat{\phi}_{2^{-j-1}L}(l,m)}{\widehat{\phi}_{2^{-j}L}(l,m)} & l < 2^{-j-1}L \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2}$$

and the high pass filter

$$\widehat{g}_j(l,m) = \begin{cases} \frac{\widehat{\psi}_{2^{-j-1}L}(l,m)}{\widehat{\phi}_{2^{-j}L}(l,m)} & l < 2^{-j-1}L \text{ and } m = 0 \\ 1 & l \geqslant 2^{-j-1}L \text{ and } m = 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.3}$$

where one of the simplest choice of ψ is $\widehat{\psi}_{2^{-j}L}(l,m) = \widehat{\phi}_{2^{-j+1}L}(l,m) - \widehat{\phi}_{2^{-j}L}(l,m)$. Clearly in this case $\widehat{g}_j = 1 - \widehat{h}_j$ The isotropic wavelet coefficients function for a function f are

$$\begin{aligned} w_j &= \phi_{2^{-j}L} * f * g_j \\ \text{or} \quad \widehat{w}_j &= \widehat{\phi}_{2^{-j}L}\widehat{f}(1-\widehat{h}_j) = \widehat{\psi}_{2^{-j}L}\widehat{f} \end{aligned} \tag{2.4}$$

One possible choice of $\widehat{\phi}_L(l,m)$ is $\frac{3}{2}B_3\left(\frac{2l}{L}\right)$, the B-spline of order 3, but in general it is not specified except for being bandlimited. However, a class of choices using radial basis functions can be found in the earlier works in [66][67], namely if we firstly ignore the bandlimit and simply set the scaling function as $\widehat{\Phi}_j(l) = \gamma(2^{-j}l)$, where $\gamma : [0,\infty) \to \mathbb{R}$ monotonously decreasing satisfies the following conditions

$$\begin{cases} \gamma \text{ continuous at zero and } \gamma(0) = 1 \\ \sum\limits_l \frac{2l+1}{4\pi}\left(\sup\limits_{x\in[l,l+1)} \left|\gamma(2^{-j}x)\right|\right)^2 < \infty \end{cases} \tag{2.5}$$

In particular the first condition implies $\lim\limits_{j\to\infty} \|f - f * \Phi_j\|_{L^2} = 0$, namely Φ_j forms an approximate identity. Examples of such functions include $(1+x)^{-s}$ with $x \in [0,\infty)$, which is natural for the Sobolev setting, the linear construction

$$\gamma(x) = \begin{cases} 1 \text{ for } x \in [0,\tau) \\ \frac{1-x}{1-\tau} \text{ for } x \in [\tau,1) \\ 0 \text{ for } x \in [1,\infty) \end{cases} \tag{2.6}$$

and the cubic construction

$$\gamma(x) = \begin{cases} (1-x)^2(1+2x) \text{ for } x \in [0,1) \\ 0 \text{ for } x \in [1,\infty) \end{cases} \tag{2.7}$$

Wavelets and its dual are chosen in this setting to meet the simple equality

$$\Psi_j * \tilde{\Psi}_j = \Phi_{j+1} * \Phi_{j+1} - \Phi_j * \Phi_j \tag{2.8}$$

and consequently there is

$$\Phi_0 * \Phi_0 + \sum_{j=0}^{\infty} \Psi_j * \Psi_j = 1 \tag{2.9}$$

Two typical choices are thus

$$\widehat{\Psi}_j = \widehat{\tilde{\Psi}}_j = \sqrt{\widehat{\Phi}_j \widehat{\Phi}_j - \widehat{\Phi}_{j+1} \widehat{\Phi}_{j+1}} \text{ and } \begin{cases} \widehat{\Psi}_j = \widehat{\Phi}_j - \widehat{\Phi}_{j+1} \\ \widehat{\tilde{\Psi}}_j = \widehat{\Phi}_j + \widehat{\Phi}_{j+1} \end{cases} \tag{2.10}$$

The advantage of this setting is that reconstruction of zonal functions follows immediately from the hierarchical property of radial basis functions without extra efforts. Indeed, by set $V_j = \{\Phi_j * \Phi_j * f : f \in L^2\}$ where $\Phi_j * \Phi_j$ is the low-pass and $W_j = \{\Psi_j * \tilde{\Psi}_j * f : f \in L^2\}$ where $\Psi_j * \tilde{\Psi}_j$ is the band-pass, a natural multi-resolution structure appears, namely

$$\begin{cases} V_0 \subset \cdots \subset V_j \subset V_{j+1} \subset \cdots \subset L^2 \\ \bigcup_{j=0}^{\infty} V_j \text{ dense in } L^2 \\ \text{If } \Phi_j * \Phi_j * f \in V_j, \text{then } \Phi_{j+1} * \Phi_{j+1} * f \in V_{j+1} \end{cases} \tag{2.11}$$

As a result any square-integrable function can be approximated by adding detailed terms from W_j level-wise. It is desirable that band-pass vanishes for lower orders, one possible way is to define γ^{L_0} equal to 1 on $[0, L_0]$ being continuous at $x = L_0$ and set

$$\widehat{\Phi}_j^{L_0}(l) = D_j^{L_0} \gamma^{L_0}(l) = \gamma^{L_0}(L_0 + 2^{-j}(l - L_0)) \tag{2.12}$$

which has the properties that $\widehat{\Phi}_j^{L_0}(l) = 1$ for $l = 0, \cdots, L_0$ and $\lim_{j \to \infty} \widehat{\Phi}_j^{L_0}(l) = 1$ for any l. Thus for arbitrarily given $f \in L^2$, $\Psi_j^{L_0} * \tilde{\Psi}_j^{L_0} * f$ is orthogonal to $\mathcal{H}_l$ for $l \leqslant L_0$. Meanwhile for bandlimited function, say $\hat{f}_l = 0$ for $l > L$, we could instead set $\widehat{\Phi}_j^{L_0}(l) = \gamma^{L_0}(2^{-j} l)$, then the corresponding wavelets have the good property that $\Psi_j^{L_0} * \tilde{\Psi}_j^{L_0} * f = 0$ whenever $2^{-j} \leqslant L_0/L$.

In the continuous setting, however, the admissibility conditions imposed on the wavelets are

$$\widehat{\Psi}_\rho(0) = 0 \text{ and } \int_0^\infty \widehat{\Psi}_\rho(l)\alpha(\rho)d\rho = 1 \text{ for } l \geqslant 1$$
$$\sum_1^\infty \frac{2l+1}{4\pi}\left(\int_a^\infty \widehat{\Psi}_\rho(l)\alpha(\rho)d\rho\right)^2 < \infty \tag{2.13}$$
$$\int_{-1}^1 \left|\int_a^\infty \Psi_\rho(t)\alpha(\rho)d\rho\right| dt < T$$

where $\alpha(\rho)d\rho$ is an arbitrary positive measure. Under the first condition there is

$$\int_a^\infty \int_{\mathbb{S}^2} \Psi_\rho(\xi\cdot\eta)f(\eta)d\eta\alpha(\rho)d\rho = \sum_l \widehat{\Phi}_a(l)\mathbb{P}_l(f) \to f \text{ as } a \to 0$$

in the sense of L^2, where $\Phi_a(t) = \int_a^\infty \Psi_\rho(t)\alpha(\rho)d\rho$ or equivalently $\widehat{\Phi}_a(l) = \int_a^\infty \widehat{\Psi}_\rho(l)\alpha(\rho)d\rho$ is well defined in L^2 since the second condition of (2.13) implies that $\Phi_a \in L^2([-1,1])$. The convergence holds because of the uniform boundedness $\left|\widehat{\Phi}_a(l)\right| \leqslant \int_{-1}^1 \left|\int_a^\infty \Psi_\rho(t)\alpha(\rho)d\rho\right| dt < T$ and the Banach-Steinhaus theorem.

To deal with the anisotropic situation, one strategy is to add an additional directional function $\mathfrak{h} \in L^2(\mathbb{S}^2)$ with $\sum_{|m|\le l} |\mathfrak{h}_{lm}|^2 \neq 0$ to form the directional wavelet

$$\hat{\psi}_{lm} = \widehat{\Psi}(l)\mathfrak{h}_{lm} \tag{2.14}$$

If $\mathfrak{h}$ is bandlimited, one can assume without loss of generality that $\mathfrak{h}$ preserves the energy on each degree l, namely $\sum_{|m|\le l} |\mathfrak{h}_{lm}|^2 = 1$, hence $\|\psi\|_2^2 = \sum_0^\infty \widehat{\Psi}(l)^2$.

The dilation operation in this formulation is defined by

$$\widehat{\mathcal{D}_2(a)\psi}_{lm} = \widehat{\Psi}(al)\mathfrak{h}_{lm} \tag{2.15}$$

By choosing a Ψ such that $supp\widehat{\Psi} \subset (a, a^{-1})$ with $a \in (0,1)$ and J the smallest integer such that $a^J L \leqslant 1$, one can define in the same way as in (2.10) that

$$\widehat{\Psi}_j(L^{-1}l) = \widehat{\Psi}(a^{-j}L^{-1}l) = \sqrt{\gamma_a^2(a^{-(j-1)}L^{-1}l) - \gamma_a^2(a^{-j}L^{-1}l)} \tag{2.16}$$

with $\gamma_a^2(t) = \frac{\int_t^1 r_a^2(u)du}{\int_a^1 r_a^2(u)du}$ and $r_a(t) = r\left(\frac{2}{1-a}(t-a)-1\right)$ for some Schwartz function r on $[-1,1]$.

Note that $supp(r_a) \subset [a, 1]$ and $\gamma_a(t) = 1$ when $t \leqslant a$, hence there is

$$supp\gamma_a(a^{-j}L^{-1}l) \subset (-\infty, a^j L]$$

and $\widehat{\psi}^j_{lm} = \widehat{\Psi}^j(l)\mathfrak{h}_{lm}$ is supported in $[a^{j+1}L, a^{j-1}L]$, which together with the scaling function

$$\widehat{\varphi}_{lm} = \gamma_a(a^{-J}L^{-1}l)\delta_{m,0} \tag{2.17}$$

give the reconstruction formula for each l within the bandlimit of $\mathfrak{h}$ that

$$\widehat{\varphi}^2_{l0} + \sum_{j=0}^{J} \sum_{|m|\leqslant l} \widehat{\psi}^2_{lm}(a^{-j}L^{-1}l) = 1 \tag{2.18}$$

in a similar manner as (2.9).

One choice for the directional function $\mathfrak{h}$, as proposed in [41], is based on the concept of steerability. The class of steerable functions on $\mathbb{R}^2$ under the steering constrain that

$$f_\theta(r, \phi) = \sum_{m=1}^{M} k_m(\theta) g_m(r, \phi) \tag{2.19}$$

for some $M \in \mathbb{N}$ and basis functions g_m were discussed firstly in [69]. In [9] this definition was applied to spherical functions with respect to the third Euler angle, a quantity that can be understood as the direction in the tangent plane.

Let us, however, generalize the definition in [41] slightly and prove a stronger result on the sphere. Denote by $\mathcal{R}(t)$ the rotation operation around axis ξ_0 by an angle t, we call a function $f \in L^2(\mathbb{S}^2)$ *steerable*, if there exists some $t_m \in \mathbb{S}^1$, $m = 1, \cdots, M$ such that

$$\mathcal{R}(t)f = \sum_{m=1}^{M} k_m(t)\mathcal{R}(t_m)f \tag{2.20}$$

holds for almost every t.

Proposition 2.1. *Steerability is equivalent to the existence of an azimuthal band limit in m for $L^2(\mathbb{S}^2)$ functions.*

Proof. In fact, if a function $f(\theta, \varphi) = \sum_{n=-\infty}^{\infty} a_n(\theta)e^{in\varphi} \in L^2(\mathbb{S}^2)$ satisfies (2.20) for some t_m and almost every $t \in [0, 2\pi]$, then

$$\sum_{n=-\infty}^{\infty} e^{in\varphi} \left[a_n(\theta) \left(e^{-int} - \sum_{m=1}^{M} k_m(t) e^{-int_m} \right) \right] = 0 \quad a.e. \tag{2.21}$$

Since Fourier series forms a Schauder basis for $L^2(\mathbb{S}^1)$, for any $a_n(\theta) \neq 0$, there is $e^{-int} - \sum_{m=1}^{M} k_m(t)e^{-int_m} = 0$ *a.e.* Without loss of generality, assume for some $L > M$ there is

$$\begin{pmatrix} e^{-i(t_1-t)} & e^{-i(t_2-t)} & \cdots & e^{-i(t_M-t)} \\ e^{-i2(t_1-t)} & e^{-i2(t_2-t)} & \cdots & e^{-i2(t_M-t)} \\ \vdots & \vdots & & \vdots \\ e^{-iL(t_1-t)} & e^{-iL(t_2-t)} & \cdots & e^{-iL(t_M-t)} \end{pmatrix} \begin{pmatrix} k_1(t) \\ k_2(t) \\ \vdots \\ k_M(t) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \tag{2.22}$$

or simply $\mathbf{G_t k_t} = \mathbf{E}$ for almost every t, which implies that rank$\mathbf{G}_t \leq M$ *a.e.*, contradictory to the linear independence of set $\{e^{ijt} : j = 1, \cdots, L\}$. Therefore the number of non-zero Fourier coefficients of f with respect to variable φ is at most M. The other direction of the proof of the equivalence is an immediate consequence of the next Lemma. $\square$

Lemma 2.2. *If $f \in L^2(\mathbb{S}^2)$ and $\hat{f}_{lm} = 0$ for any $|m| \geqslant L + 1 \in \mathbb{N}_0$, then for any $t \in [0, 2\pi]$,*

$$\sum_{n=-L}^{L} k(t - t_n)\mathcal{R}(t_n)f = \mathcal{R}(t)f \tag{2.23}$$

with $t_n = \frac{2\pi n}{2L+1}$ and $k \in L^2(\mathbb{S}^1)$ given by $k(t) = \sum_{n=-L}^{L} \frac{1}{2L+1} e^{int}$.

Proof. Since

$$\frac{1}{2L+1} \sum_{|n| \leqslant L} e^{-imt_n} = \begin{cases} 1 & \text{for} \quad m = 0 \\ 0 & \text{for} \quad m \neq 0 \end{cases} \tag{2.24}$$

the Fourier coefficients of both sides of (2.23) coincide, namely

$$\begin{aligned} (\widehat{\mathcal{R}(t)f})_{lm'} &= e^{im't}\hat{f}_{lm'} \\ &= \frac{1}{2L+1} \sum_{|m| \leqslant L} \sum_{|n| \leqslant L} e^{i(m'-n)t_m} e^{int}\hat{f}_{lm'} \\ &= \frac{1}{2L+1} \sum_{|m| \leqslant L} \sum_{|n| \leqslant L} e^{in(t-t_m)} (\widehat{\mathcal{R}(t_m)f})_{lm'} \\ &= \sum_{|m| \leqslant L} k(t - t_m)(\widehat{\mathcal{R}(t_m)f})_{lm'} \end{aligned}$$

Therefore the equality (2.23) holds.

$\square$

Under the assumption that $\mathfrak{h}$ is steerable, namely $\mathcal{R}(t)\mathfrak{h} = \sum_{|m'|\leqslant L} k_{m'}(t)\mathcal{R}(t_{m'})\mathfrak{h}$ for some finite L, there is

$$\begin{aligned}
\mathcal{R}(t)\psi &= \sum_{l} \widehat{\Psi}(l)\mathcal{R}(t)\mathbb{P}_l\mathfrak{h} \\
&= \sum_{|m'|\leqslant L} k_{m'}(t) \sum_{l} \widehat{\Psi}(l)\mathbb{P}_l\left(\mathcal{R}(t_{m'})\mathfrak{h}\right) \\
&= \sum_{|m'|\leqslant L} k_{m'}(t)\mathcal{R}(t_{m'})\psi
\end{aligned} \tag{2.25}$$

namely ψ is steerable too. In this case there is

$$\begin{aligned}
\langle \mathcal{R}(t)\psi^j, \mathcal{R}(t^{'})\psi^j\rangle_{L^2} &= \sum_{l} \sum_{|m'|\leqslant \min\{l,L\}} e^{-im^{'}(t-t^{'})} \left|\hat{\psi}^j_{lm'}\right|^2 \\
&= \sum_{l} \Delta_l(t-t^{'}) \left|\hat{\Psi}_j(l)\right|^2
\end{aligned} \tag{2.26}$$

where $\Delta_l(t) = \sum_{|m'|\leqslant L} |\mathfrak{h}_{lm'}|^2 e^{-im^{'}(t)}$. Intuitively speaking, if we choose $0 = t^1 < t^2 < \cdots < t^N < 2\pi$ and form a $N \times N$ matrix

$$M = \langle \mathcal{R}(t^s)\psi^j, \mathcal{R}(t^u)\psi^j\rangle_{L^2} \tag{2.27}$$

with $1 \leqslant s, u \leqslant N$, the diagonal elements are 1's, while the smaller the off diagonal elements are, the more directional ψ^j is.

Within this approach, the best known work in frame properties aspect is probably [30], where a localized tight frame is constructed in the following delicate way. By taking a continuous function $\mathfrak{a}$ supported in $[\frac{1}{2}, 2]$, for instance $\mathfrak{a}(t) = m_0(\pi \log_2(t))$ with m_0 the standard orthogonal wavelet mask on real line, such that

$$|\mathfrak{a}(t)|^2 + |\mathfrak{a}(2t)|^2 = 1 \text{ on } [\frac{1}{2}, 1] \tag{2.28}$$

We have obviously for any $J \in \mathbb{Z}_+$ that

$$\begin{aligned}
\mathfrak{b}(2^{-J}t) &:= \sum_{j=-\infty}^{J} \left|\mathfrak{a}(2^{-j}t)\right|^2 \\
&= \begin{cases} \left|\mathfrak{a}(2^{-j}t)\right|^2 + \left|\mathfrak{a}(2^{-j+1}t)\right|^2 = 1 & \text{for } t \in 2^j[\frac{1}{2}, 1], j \leqslant J \\ \left|\mathfrak{a}(2^{-J}t)\right|^2 & \text{for } t \in 2^{J+1}[\frac{1}{2}, 1] \end{cases}
\end{aligned} \tag{2.29}$$

In particular, when $J \to \infty$, $\sum\limits_{j=-\infty}^{\infty} \left|\mathfrak{a}(2^{-j}t)\right|^2 = 1$ on $(0,\infty)$. Thus, if we define for $j \leqslant J$ that

$$A_j^d(t) = \begin{cases} \frac{1}{\pi}\sum\limits_{l=1}^{\infty} \mathfrak{a}(2^{-j}l)\cos(l\arccos t) & \text{for } d=2 \\ \sum\limits_{l=0}^{\infty} \mathfrak{a}(2^{-(j+j_d)}(l+\lambda_d))Z_l^d(t) & \text{for } d \geqslant 3 \end{cases} \tag{2.30}$$

and

$$B_J^d(t) = \begin{cases} \frac{1}{2\pi} + \frac{1}{\pi}\sum\limits_{l=1}^{\infty} \mathfrak{b}(2^{-J}l)\cos(l\arccos t) & \text{for } d=2 \\ \sum\limits_{l=0}^{\infty} \mathfrak{b}(2^{-(J+j_d)}(l+\lambda_d))Z_l^d(t) & \text{for } d \geqslant 3 \end{cases} \tag{2.31}$$

where $j_d = \log_2[\lambda_d]$ for $d \geqslant 3$ and $j_1 = 0$ are so that the minimal eigenvalue of the operator

$$\mathcal{L}_n = \sqrt{\lambda_d - \Delta_{\mathbb{S}^{d-1}}} = \sum_{l=0}^{\infty}(l+\lambda_d)\mathbb{P}_l^d \tag{2.32}$$

lies in interval $[1,2]$, then for any $f \in L^2(\mathbb{S}^{d-1})$

$$\sum_{j=-\infty}^{J} \overline{A_j^d} * A_j^d * f = \begin{cases} (B_J^d - P_0) * f & \text{for } d=2 \\ B_J^d * f & \text{for } d \geqslant 3 \end{cases} \tag{2.33}$$

and by Funk-Hecke formula and the support of $\mathfrak{a}$

$$\begin{aligned} \langle \overline{A_j^d} * A_{j'}^d * f, f\rangle &= \langle \sum_l \overline{\mathfrak{a}(2^{-(j+j_d)}(l+\lambda_d))}\mathfrak{a}(2^{-(j'+j_d)}(l+\lambda_d))\mathbb{P}_l^d f, f\rangle \\ &= 0 \end{aligned} \tag{2.34}$$

for any $|j-j'| \geqslant 2$, $d \geqslant 3$.

Suppose $\mathbb{S}^{d-1}$ has a subdivision into $\{\Omega_i\}_{i\in I}$ labeled by a set of points $\mathcal{V} = \{p_i\}_{i\in I}$. For a mesh with good uniformity, namely the *mesh ratio*

$$\rho_{\mathcal{V}} = \frac{h_{\mathcal{V}}}{q_{\mathcal{V}}} = \frac{\sup\limits_{x\in\mathbb{S}^{d-1}} \inf\limits_{p_i\in\mathcal{V}} d(x,p_i)}{\frac{1}{2}\min\limits_{p_i\neq p_j} d(p_i,p_j)} \tag{2.35}$$

is larger than 2, where $h_{\mathcal{V}}$ is called *mesh norm* and $q_{\mathcal{V}}$ called *separation radius* of $\mathcal{V}$ respectively, there exists a nested sequence $\mathcal{V}_k \subset \mathcal{V}_{k+1}$ such that

$$\frac{1}{4}h_{\mathcal{V}_k} < h_{\mathcal{V}_{k+1}} < \frac{1}{2}h_{\mathcal{V}_k} \tag{2.36}$$

It is proved in [30] as main theorems, here we cite as lemmas without proof, that

Lemma 2.3. *There exist constants s_d depending solely on dimension d such that, for any $\delta \in (0, \frac{1}{2})$ and integer $L \in \mathbb{Z}_+$, whenever $\|\mathcal{V}\| = \max_{p \in \mathcal{V}} diam(\Omega_i) \leqslant \delta s_d^{-1}(L + \lambda_n)^{-1}$,*

$$(1-\delta)\|f\|_1 \leqslant \sum_{p \in \mathcal{V}} |f(p)||\Omega_p| \leqslant (1+\delta)\|f\|_1 \tag{2.37}$$

holds for spherical harmonics in $\Pi_L^d = \bigoplus_{l \leqslant L} \mathcal{H}_l^d$; there exist positive weights $w(p)$ such that

$$\int_{\mathbb{S}^{d-1}} f(\eta) d\eta = \sum_{p \in \mathcal{V}} w(p) f(p) \tag{2.38}$$

and $w(p)$ have the following bounds

$$\frac{1-2\delta}{1-\delta} |\Omega_p| \leqslant w(p) \leqslant \frac{|\mathbb{S}^{d-1}|}{dim(\Pi_{[L/2]}^d)} \tag{2.39}$$

Lemma 2.4. *For $\mathfrak{a} \in C^k(\mathbb{R})$ with $k > \max\{d-1, 2\}$, if $f \in L^q(\mathbb{S}^{d-1})$ with $1 \leqslant q \leqslant \infty$, then there exists constant $C_{b,k,d}$ such that*

$$\|f - B_J^d * f\|_q \leqslant C_{b,k,d} dist_{L^q}(f, \Pi_L^d) \tag{2.40}$$

As a result, on $j-$th level mesh $\mathcal{V}_j$, by defining

$$\psi_{j,p}(\eta) = \sqrt{w_{j,p}} A_j^d(\eta \cdot p) \tag{2.41}$$

with $\mathfrak{a} \in C^k$, we have for $f \in C(\mathbb{S}^{d-1})$ or $f \in L^q(\mathbb{S}^{d-1})$ with $1 \leqslant q < \infty$,

$$f = \sum_{j=0}^{\infty} \sum_{p \in \mathcal{V}_j} \langle f, \psi_{j,p} \rangle \psi_{j,p} \tag{2.42}$$

with convergence in the corresponding space norms. For $f \in L^2$, it is equivalent to

$$\|f\|^2 = \begin{cases} \frac{1}{2\pi} |\langle f, 1 \rangle|^2 + \sum\limits_{j=0}^{\infty} \sum\limits_{p \in \mathcal{V}_j} |\langle f, \psi_{j,p} \rangle|^2 & \text{for } d = 2 \\ \sum\limits_{j=0}^{\infty} \sum\limits_{p \in \mathcal{V}_j} |\langle f, \psi_{j,p} \rangle|^2 & \text{for } d \geqslant 3 \end{cases} \tag{2.43}$$

namely $\{\psi_{j,p}\}$ forms a tight frame. Meanwhile, when $2^{-j} < \epsilon$, there is the integral expression of

$$A_\epsilon^d(t) = \begin{cases} \frac{1}{\pi}\sum\limits_{l=1}^{\infty} \mathfrak{a}(\epsilon)\cos(l\arccos t) & \text{for } d=2 \\ \sum\limits_{l=0}^{\infty} \mathfrak{a}(\epsilon(l+\lambda_d))Z_l^d(t) & \text{for } d \geqslant 3 \end{cases}$$

that

$$A_\epsilon^d(\cos\theta) = \frac{\gamma_d}{\sin^{d-1}\theta}\int_\theta^\pi \frac{C_{\epsilon,d}(\phi)}{(\cos\theta - \cos\phi)}d\phi \tag{2.44}$$

where $\gamma_d = \frac{2^{\lambda_d}\Gamma(\lambda_d+1/2)}{\sqrt{\pi}|\mathbb{S}^{d-1}|\Gamma(\lambda_d)\Gamma(2\lambda_d)}$ and

$$C_{\epsilon,d}(\phi) = \frac{1}{2\epsilon}\sum_{n\in\mathbb{Z}}(-1)^{(d-2)n}\tilde{Q}_{d-2}(i\frac{d}{d\phi})\hat{\mathfrak{a}}(\frac{\phi+w\pi n}{\epsilon}) \tag{2.45}$$

with

$$\tilde{Q}_{d-2}(z) = \begin{cases} \prod\limits_{j=1}^{[\frac{d-2}{2}]} \left(z^2 - (\lambda_d - j)^2\right) z\sin(\lambda_d\pi) & \text{if } j \text{ even} \\ \prod\limits_{j=1}^{[\frac{d-2}{2}]} \left(z^2 - (\lambda_d - j)^2\right) \cos(\lambda_d\pi) & \text{if } j \text{ odd} \end{cases} \tag{2.46}$$

It leads to an estimation for $\mathfrak{a} \in C^k(\mathbb{R})$ that

$$|A_\epsilon^d(\cos\theta)| \leqslant \frac{\beta_{d,k,a}\epsilon^{-d+1}}{1+(\frac{\theta}{\epsilon})^k} \tag{2.47}$$

with some constants $\beta_{d,k,a}$. Thus we see that when j and k increase, $|\psi_{j,p}|$ decrease in scale value as $2^{-(j+j_d)k}$, hence is localized.

2.2 Dilation through stereographic projection

In contrast to doing dilation in frequency domain, this section we discuss dilation through geometric approach. As we have encountered at the beginning of the chapter and we shall see in the review part later, one common strategy is to wisely cover the whole manifold by local patches, and then construct a dictionary as it was done on the plane followed by lifting it back. This method has the advantage that it can be applied to any manifolds and it has localized nature.

On the sphere there is much more to say than on a general manifold, not only because the stereographic projection is obviously ideal for building a one-to-one correspondence with the plane, but also that the Euclidean sphere has exquisite group structures that has been introduced above. Stereographic map is conformal and can be used to generate the whole system globally, while by choosing localized generators it allows us to construct frame systems with ideal local properties. Although at the pole it may cause distortion it avoids the consistency problem arising from different patches, hence much easier to implement in my opinion. The first work in this approach probably dates back to [92][93], where admissibility condition is proposed without successfully formulating frames.

In this section, however, dilation is done in both isotropic and parabolic manner. Especially the second brings much difference when one deals with anisotropic problems, exactly for this reason I shall call the corresponding generators "spherical shearlets" or "spherical α-wavelets". A long lasting unsolved problem in the stereographic projection approach before the writing of this dissertation is the lack of a constructible tight frame. Discussion about that is going to be delayed till the next chapter, where exact design of a class of tight frames is given.

By identifying a 2-sphere with the homogeneous space $SO(3)/SO(2)$, the sphere is embedded into the rotation group and group operations can be applied naturally. Therefore the translation on the sphere can be achieved by rotation. However, in comparison with translation in the Euclidean plane, it obviously brings much more complicated work since parameters are located in the rotation group instead of $\mathbb{R}$. Let $d\Omega = \sin\theta d\theta d\varphi$ be the rotation invariant measure on the unit sphere, and $d\mathbf{x}$ be the Lebesgue measure on $\mathbb{R}^2$. Denote by σ an element in the group $SO(3)$ and by $d\sigma$ the Haar measure on it. The Haar measure is normalized so that the whole group $SO(3)$ has volume one. Recall that the left regular unitary representation defined by

$$(\mathcal{R}(\sigma)f)(\omega) = f(\sigma^{-1}\omega)$$

where $\sigma \in SO(3)$. Related to the rotation operator $\mathcal{R}(\sigma)$ there is the Wigner D-matrix which consists of coefficients with respect to the normalized orthogonal spherical harmonics. Peter-Weyl Theorem tells us that these coefficients are dense in $L^2(\mathbb{S}^2)$ and $\mathcal{R}$ can be written as the direct sum of finite-dimensional irreducible representations. $L^2(\mathbb{S}^2)$ is correspondingly decomposed into $\mathcal{R}$-invariant vector subspaces, those are exactly the eigenspaces $\mathcal{H}_l$ corresponding to the different eigenvalues of the Laplace-Beltrami operator on the sphere, as we have seen in the introduction.

Schulr's orthogonality relations in this case are

$$\int_{SO(3)} D^l_{mn}(\sigma)\overline{D^{l'}_{m'n'}(\sigma)}d\sigma = \frac{1}{d_l}\delta_{ll'}\delta_{mm'}\delta_{nn'} \tag{2.48}$$

where indexes m,n and m', n' are integers of absolute value no larger than l and l' respectively.

Let $\omega = (\theta, \varphi)$ denote the polar coordinate of a point on the unit sphere. In particular, $\theta = \pi$ is the north pole ξ_0. The stereographic projection $\pi : \mathbb{S}^2 \backslash \{\xi_0\} \mapsto \mathbb{R}^2$, by

$$\pi(\theta, \varphi) = (r\cos\varphi, r\sin\varphi) \tag{2.49}$$

with $r = 2\tan(\frac{\theta}{2})$, gives an isomorphism, and its inverse is denoted by π^{-1}.

One strategy of doing dilation on the sphere is to utilize the dilation operator $d(a)$ on $L^2(\mathbb{R}^2, d\mathbf{x})$ defined as

$$(d(a)h)(x, y) = a^{-(1+\alpha)/2}h(a^{-1}x, a^{-\alpha}y) \tag{2.50}$$

with $\alpha \in [0, 1]$.

Definition 2.5. Given $a \in \mathbb{R}_+$ and $f \in L^2(\mathbb{S}^2, d\Omega)$, define $\mathcal{D}(a) : L^2(\mathbb{S}^2, d\Omega) \mapsto L^2(\mathbb{S}^2, d\Omega)$ as

$$\mathcal{D}(a)f(\omega) = \begin{cases} U^{-1}d(a)Uf(\omega) & \omega \neq \xi_0 \\ a^{\frac{-1-\alpha}{2}}f(\xi_0) & \omega = \xi_0 \end{cases}$$

where U is the operator such that

$$(Uf)(\mathbf{x}) = \nu(\pi^{-1}\mathbf{x})f(\pi^{-1}\mathbf{x})$$

with $\nu(\theta, \varphi) = \cos^2(\theta/2)$ guaranteeing the unitarity of the operator.

Let us make the convention that throughout this chapter α is chosen as a positive number no larger than 1. In the special case $\alpha = \frac{1}{2}$, $\mathcal{D}(a)$ will be called *parabolic spherical dilation operator*.

A simple calculation shows that $\omega = (\theta, \varphi)$ and the point after dilation $\omega_{\frac{1}{a}} = \hat{d}(a)(\theta, \varphi) = (\theta_{\frac{1}{a}}, \varphi_{\frac{1}{a}})$, are related by

$$\begin{cases} \tan(\theta_{\frac{1}{a}}/2) = \sqrt{\gamma(a, \varphi)}\tan(\theta/2) \\ \tan\varphi_{\frac{1}{a}} = a^{1-\alpha}\tan\varphi \quad \text{when } \varphi \neq \frac{\pi}{2} \text{ and } \varphi \neq \frac{3\pi}{2} \\ \varphi_{\frac{1}{a}} = \varphi \quad \text{when } \varphi = \frac{\pi}{2} \text{ or } \frac{3\pi}{2} \end{cases} \tag{2.51}$$

where $\sqrt{\gamma(a, \varphi)} = (a^{-2}\cos^2\varphi + a^{-2\alpha}\sin^2\varphi)^{1/2}$.

Let $\theta^{'} = \theta_{1/a}$, $\varphi^{'} = \varphi_{1/a}$. By (2.51), it is easy to see the following relations that are going to facilitate our mathematical deduction.

$$\begin{aligned}
\sin^2\varphi &= \frac{a^{2\alpha-2}\tan^2\varphi^{'}}{1+a^{2\alpha-2}\tan^2\varphi^{'}}, \quad \cos^2\varphi = \frac{1}{1+a^{2\alpha-2}\tan^2\varphi^{'}} \\
\sin\theta &= \frac{2\tan(\frac{\theta^{'}}{2})\cos\varphi^{'}\sqrt{a^2+a^{2\alpha}\tan^2\varphi^{'}}}{1+\tan^2(\frac{\theta^{'}}{2})\cos^2\varphi^{'}(a^2+a^{2\alpha}\tan^2\varphi^{'})} \\
\cos\theta &= \frac{1-\tan^2(\frac{\theta^{'}}{2})\cos^2\varphi^{'}(a^2+a^{2\alpha}\tan^2\varphi^{'})}{1+\tan^2(\frac{\theta^{'}}{2})\cos^2\varphi^{'}(a^2+a^{2\alpha}\tan^2\varphi^{'})}
\end{aligned} \tag{2.52}$$

With these, we arrive at an explicit formulation of the spherical dilation operator.

Proposition 2.6. *For* $f \in L^2(\mathbb{S}^2, d\Omega)$,

$$\|\mathcal{D}(a)f\|_2 = \|f\|_2 \tag{2.53}$$

$$(\mathcal{D}(a)f)(\omega) = \begin{cases} \sqrt{\lambda(a,\theta,\varphi)}f(\omega_{\frac{1}{a}}) & \omega \neq \xi_0 \\ a^{\frac{-1-\alpha}{2}}f(\xi_0) & \omega = \xi_0 \end{cases} \tag{2.54}$$

where

$$\sqrt{\lambda(a,\theta,\varphi)} = \frac{2a^{\frac{3-\alpha}{2}}}{\Phi_a^-\cos\theta+\Phi_a^+} = a^{-(1+\alpha)/2}(1+J^2)\cos^2(\frac{\theta^{'}}{2}) \tag{2.55}$$

$\Phi_a^{\pm} = a^2[1\pm\gamma(a,\varphi)]$ *and* $J = \tan\frac{\theta}{2} = \tan(\frac{\theta^{'}}{2})\cos\varphi^{'}\sqrt{(a^2+a^{2\alpha}\tan^2\varphi^{'})}$.

Proof. $\|\mathcal{D}(a)f\|_2 = \|f\|_2$ comes from the fact that both $d(a)$ and U are unitary operators. From the definition (2.5) it follows immediately that

$$U^{-1}g(\theta,\varphi) = \frac{1}{\nu(\theta,\varphi)}g(\pi(\theta,\varphi)) \tag{2.56}$$

and from (2.49) that

$$\pi^{-1}(x_1,x_2) = (2\arctan(\frac{\sqrt{x_1^2+x_2^2}}{2}), \arctan(\frac{x_2}{x_1})) \tag{2.57}$$

for any $g \in L^2(\mathbb{S}^2, d\Omega)$ and $\mathbf{x} = (x_1, x_2)$ on the plane. Therefore by the definitions (2.50) and (2.5) for $\theta \neq \pi$

$$\begin{aligned}
\mathcal{D}(a)f(\theta,\varphi) &= [U^{-1}d(a)Uf](\theta,\varphi) \\
&= \frac{1}{\nu(\theta,\varphi)}(d(a)Uf)(\pi(\theta,\varphi)) \\
&= \frac{a^{-(1+\alpha)/2}}{\nu(\theta,\varphi)}\nu(\pi^{-1}(\frac{2\tan\frac{\theta}{2}\cos\varphi}{a}, \frac{2\tan\frac{\theta}{2}\sin\varphi}{a^\alpha}))f(\omega_{1/a}) \\
&= \frac{a^{-(1+\alpha)/2}}{\cos^2\frac{\theta}{2}}\cos^2(\arctan[(\frac{\cos^2\varphi}{a^2} + \frac{\sin^2\varphi}{a^{2\alpha}})^{1/2}\tan\frac{\theta}{2}])f(\omega_{1/a}) \\
&= \frac{2a^{(3-\alpha)/2}f(\omega_{1/a})}{(a^2 - a^{2-2\alpha}\sin^2\varphi - \cos^2\varphi)\cos\theta + a^2 + a^{2-2\alpha}\sin^2\varphi + \cos^2\varphi}
\end{aligned}$$

which is exactly the expression (2.54).

Meanwhile, from (2.52) we deduce that

$$\Phi_a^{\pm} = a^2 \pm \frac{1}{\cos^2\varphi'(1 + a^{2\alpha-2}\tan^2\varphi')}$$

hence

$$\begin{aligned}
\Phi_a^- \cos\theta + \Phi_a^+ &= a^2(\cos\theta + 1) + \frac{a^2\tan^2(\frac{\theta'}{2})(1-\cos\theta)}{J^2} \\
&= \frac{2a^2}{1+J^2} + \frac{2a^2\tan^2(\frac{\theta'}{2})}{1+J^2} \\
&= \frac{2a^2\sec^2(\frac{\theta'}{2})}{1+J^2}
\end{aligned}$$

and (2.55) follows.

□

Remark 2.7. The coefficient $\lambda(a,\theta,\varphi)$ can be alternatively defined as the Radon-Nikodym derivative $\frac{d\Omega(\omega_{1/a})}{d\Omega(\omega)}$, and it is easy to check these two different ways of definition give the same result. In this sense, λ can be interpreted as the change of measure caused by dilation operation.

With the rotation operator as well as the dilation operator, the continuous spherical wavelets on the sphere is generated by a single function.

Definition 2.8. For $\psi \in L^2(\mathbb{S}^2, d\Omega)$, define *spherical α-wavelet system* on $\mathbb{S}^2$ as

$$\{\psi_{\sigma,a}(\omega) = \mathcal{R}(\sigma)\mathcal{D}(a)\psi(\omega) = \sqrt{\lambda(a,\theta,\varphi)}\psi((\sigma^{-1}\omega)_{1/a}) : \sigma \in SO(3), a \in \mathbb{R}^+\}$$

in particular when $\alpha \neq 1$, let us call it a continuous *spherical shearlet system.* Define *spherical α-wavelets/shearlet transform* as

$$\mathcal{S}_f(\sigma, a) = \langle f, \psi_{\sigma,a}\rangle.$$

It is natural to ask if we can exchange the order of the rotation and the dilation as defined above, namely if $\mathcal{D}(a)$ is a rotation commutative operator, unfortunately the answer is no in general. In fact, there are the following simple counter examples.

Example 2.9.

(i) Let $[\mathcal{D}(a), \mathcal{R}(\sigma)] = \mathcal{D}(a)\mathcal{R}(\sigma) - \mathcal{R}(\sigma)\mathcal{D}(a)$.

$$\sigma_3(\varphi) = \begin{pmatrix} \cos\varphi & \sin\varphi & 0 \\ -\sin\varphi & \cos\varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}, \ \alpha < 1 \text{ in (2.51), then as } a \to 0,$$

$$[\mathcal{D}(a), \mathcal{R}(\sigma_3(\varphi))]f(\theta_0, \varphi_0) \longrightarrow f(\pi, \varphi) - f(\pi, 0)$$

or $f(\pi, \pi + \varphi) - f(\pi, \pi)$.

(ii) $$\sigma_2(\theta) = \begin{pmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{pmatrix}, \ \alpha < 1, \text{ then as } a \to 0,$$

$$[\mathcal{D}(a), \mathcal{R}(\sigma_2(\theta))]f(\theta_0, \varphi_0) \longrightarrow f(\pi + \theta, 0) - f(\pi, 0)$$

or $f(\pi + \theta, \pi) - f(\pi, \pi)$, depending on whether $\tan\varphi \geqslant 0$ or $\tan\varphi < 0$.

2.3 Spaces of admissible α-wavelets/shearlets

Like the Fourier basis in $\mathbb{R}^n$, spherical harmonics have the draw back that they are not sensitive to local behavior in the spatial domain, or more precisely, a perturbation of the function value at a point may lead to the change of all coefficients and we are forced to do integration on whole sphere. Consequently they are insufficient in representing functions of high frequency. This motivates us to construct localized generators.

Definition 2.10. A spherical α-wavelet ψ is *admissible*, if

$$0 < C_\psi^l = \frac{1}{2l+1} \sum_{|m| \le l} \int_0^\infty \frac{da}{a^3} |(\widehat{\psi}_a)_{lm}|^2 < \infty$$

where $(\widehat{\psi}_a)_{lm}$ is the Fourier coefficients of the dilated function $\psi_a = \mathcal{D}(a)\psi$ with respect to spherical harmonics. Denote by $\mathcal{A}$ the set of admissible functions.

Without loss of generality we use spherical shearlet for the development of the theory in the rest of this section.

Proposition 2.11. *When a spherical shearlet is admissible, the following reconstruction formula holds*

$$\int_0^\infty \frac{da}{a^3} \int_{SO(3)} \mathcal{S}_f(\sigma, a) \psi_{\sigma,a}(\omega) d\sigma = \sum_l \sum_{|m| \le l} C_\psi^l \widehat{f_{lm}} Y_l^m(\omega).$$

Proof. By definition

$$D_{mn}^l(\sigma) = \langle \mathcal{R}(\sigma) Y_l^m, Y_l^n \rangle \tag{2.58}$$

Since $\mathcal{H}_l$ are invariant subspaces under $\mathcal{R}(\sigma)$, it follows immediately that

$$[\widehat{\mathcal{R}(\sigma) f}]_{lm} = \sum_{|n| \le l} D_{mn}^l(\sigma) \widehat{f}_{ln} \tag{2.59}$$

Therefore (2.48) and (2.59) together give us

$$\begin{aligned}
&\int_0^\infty \frac{da}{a^3} \int_{SO(3)} \langle f, \mathcal{R}(\sigma)\psi_a \rangle [\mathcal{R}(\sigma)\psi_a](\omega) \\
&= \int_0^\infty \frac{da}{a^3} \int_{SO(3)} \langle \sum_l \sum_{|m| \le l} \widehat{f_{lm}} Y_l^m, \mathcal{R}(\sigma)\psi_a \rangle \sum_{l'} \sum_{|n| \le l'} [\widehat{\mathcal{R}(\sigma)\psi_a}]_{l',n} Y_{l'}^n(\omega) \\
&= \int_0^\infty \frac{da}{a^3} \int_{SO(3)} \sum_{l,l'} \sum_{|m| \le l} \widehat{f_{lm}} \sum_{|n| \le l} \overline{D_{mn}^l}(\sigma) \overline{\widehat{(\psi_a)}_{ln}} \sum_{|m'| \le l'} \sum_{|n'| \le l'} \\
&\quad D_{m'n'}^{l'}(\sigma) \widehat{(\psi_a)}_{l'n'} Y_{l'}^{m'}(\omega) \\
&= \sum_l \frac{1}{2l+1} \int_0^\infty \frac{da}{a^3} \sum_{|m| \le l} \sum_{|n| \le l} |\widehat{(\psi_a)}_{ln}|^2 \widehat{f_{lm}} Y_l^m(\omega) \\
&= \sum_l \sum_{|m| \le l} C_\psi^l \widehat{f_{lm}} Y_l^m(\omega)
\end{aligned}$$

Thus the reconstruction formula holds when C_ψ^l in Definition 2.10 has finite value for every l. $\square$

In particular, if C_ψ^l is positive and independent of l, identity in Proposition 2.11 becomes

$$f(\omega) = C\int_0^\infty \frac{da}{a^3}\int_{SO(3)} \mathcal{S}_f(\sigma, a)\psi_{\sigma,a}(\omega)d\sigma \tag{2.60}$$

for some constant $C > 0$. The question concerning the existence of such construction arises, but an affirmative answer comes after much efforts.

Furthermore, the observation

$$\sum_{|m|\le l} |(\widehat{\psi}_a)_{lm}|_2^2 = \|\mathbb{P}_l\psi_a\|_2^2 \leqslant \|\psi_a\|_2^2 = \|\psi\|_2^2 < \infty \tag{2.61}$$

and

$$\sum_l \|\mathbb{P}_l\psi\|_2^2 = \|\psi\|_2^2 \tag{2.62}$$

show that $C_\psi^l = 0$ for all l only if ψ vanishes identically, and that under the assumption

$$C_\psi^l = \frac{1}{2l+1}\int_0^\infty \frac{\|\mathbb{P}_l\psi_a\|_2^2}{a^3}da < \infty$$

the space of admissible functions is closed under certain algebraic rules.

Lemma 2.12. *For $\psi \in L^2(\mathbb{S}^2)$, the two conditions:*
(i) $\psi \in \mathcal{A}$
(ii) $\|\mathbb{P}_l\psi_a\|_2 = o(a)$ *as* $a \to 0$ *for* $l \in \mathbb{Z}_+$
are equivalent. In particular, it indicates that

(i′) $f_1,\ f_2 \in \mathcal{A} \Rightarrow c_1f_1 + c_2f_2 \in \mathcal{A}$ *for any pair* $c_1^2 + c_2^2 > 0$

(ii′) If $\frac{1}{p} + \frac{1}{q} = 1$ *and* $p, q > 0,\ f_1,\ f_2 \in \mathcal{A} \Rightarrow f_1^{\frac{1}{p}} f_2^{\frac{1}{q}} \in \mathcal{A}$

(iii′) $\{f_n\} \subset \mathcal{A},\ f_n \rightrightarrows f \Rightarrow f \in \mathcal{A}$

Let us define some important function spaces that will appear frequently through out this section and later.

Definition 2.13. Let $\mathcal{B}_n$ be the subset of square integrable functions with the property that $\lim\limits_{\theta\to\pi} |f(\theta, \varphi) \cdot \tan^n \frac{\theta}{2}|$ exists and being bounded. Define

$$\mathcal{N} = \{f \in L^2(\mathbb{S}^2) : \int_0^{2\pi}\int_0^\pi f(\theta, \varphi)\tan(\frac{\theta}{2})d\theta d\varphi = 0\}$$

and for $\xi \in \mathbb{S}^2$, let

$$\mathcal{K}_\Theta(\xi) = \{f \in L^\infty(\mathbb{S}^2) : f(\xi \cdot \eta) = 0 \;\; a.e. \text{ for } \xi \cdot \eta \geqslant \cos\Theta\}.$$

In particular, f is called a *hollow pole function*, if $f \in \mathcal{K}(\xi_0) = \bigcup_\Theta \mathcal{K}_\Theta(\xi_0)$.

By definition it is clear that $\mathcal{K}_\Theta \subset \mathcal{B}_n$, hence we have the inclusion relation $\mathcal{B}_n \supset \mathcal{K}$. Furthermore, there is $\overline{\mathcal{K}} = \mathcal{B}_n$, where the closure can be taken both in L^∞ and L^2 norm.

Let us take the following formula from [60]

$$\begin{aligned} P_l(\cos\theta) = & \left(\frac{\theta}{\sin\theta}\right)^{\frac{1}{2}} J_0\left((l+\frac{1}{2})\theta\right) \\ & + \begin{cases} \theta^{1/2} O(l^{\frac{-3}{2}}) & \text{if} \quad c/l \leqslant \theta \leqslant \pi - \epsilon \\ \theta^2 O(1) & \text{if} \quad 0 < \theta \leqslant c/l \end{cases} \end{aligned} \tag{2.63}$$

where J_ν are the Bessel functions (4.52), but here we need its integral form

$$J_\nu(z) = \frac{(z/2)^\nu}{\Gamma(\nu+\frac{1}{2})\Gamma(\frac{1}{2})} \int_{-1}^{1} (1-t^2)^{\nu-\frac{1}{2}} e^{izt} dt \tag{2.64}$$

Taking derivative leads to the following asymptotic result.

Lemma 2.14.

$$\begin{aligned} P_l^{(m)}(\cos\theta) = & \frac{\Gamma(l+m+1)}{(l+\frac{1}{2})^m (l-m)! \sin^m\theta} \left(\frac{\theta}{\sin\theta}\right)^{\frac{1}{2}} J_m\left((l+\frac{1}{2})\theta\right) \\ & + \begin{cases} \theta^{1/2} O(l^{\frac{-3}{2}}) & \textit{if} \quad c/l \leqslant \theta \leqslant \pi - \epsilon \\ \theta^{m+2} O(l^m) & \textit{if} \quad 0 < \theta \leqslant c/l \end{cases} \end{aligned}$$

With the preparation above we are standing at the point to prove the main result of this section. It indicates how to construct a shearlet system meeting the admissibility condition. Without loss of generality let us take $\alpha = \frac{1}{2}$, the parabolic case, for the simplicity of the proof.

Theorem 2.15. *$\mathcal{B}_n \cap \mathcal{N} = \mathcal{B}_n \cap \mathcal{A}$ for $n \geqslant 3$.*

Proof. We firstly prove that $\mathcal{K} \cap \mathcal{N} = \mathcal{K} \cap \mathcal{A}$. Set $\theta^{'} = \theta_{1/a}$, $\varphi^{'} = \varphi_{1/a}$, $d\Omega^{'} = \sin\theta^{'} d\theta^{'} d\varphi^{'}$ and suppose $|\psi(\theta,\varphi) \cdot \tan\frac{\theta}{2}| \leqslant M$ *a.e.*. By Lemma 2.12, in order to prove mutual inclusion we only have to check that $\psi \in \mathcal{N}$ if and only if $\|\mathbb{P}_l \psi_a\|_2^2 = \sum_{|m| \leq l} |(\widehat{\psi_a})_{lm}|^2 = o(a^2)$.

By (1.15),

$$Y_l^m(\theta,\varphi) = c_{l,m} e^{im\varphi} \sin^m(\theta) P_l^{(m)}(\cos\theta)$$

with $c_{l,m} = (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}}$.

Consider $\psi_\Theta = \chi_{[0,2\pi]\times[0,\Theta]}\psi \in \mathcal{K}$. Using the fact that $\lambda(a,\theta,\varphi) = \frac{d\Omega'}{d\Omega}$, we get

$$\begin{aligned}
(\widehat{(\psi_\Theta)_a})_{lm} &= \int_{\mathbb{S}^2} \overline{Y_l^m(\omega)} (\mathcal{D}(a)\psi_\Theta)(\omega) d\Omega \\
&= \int_{\mathbb{S}^2} \overline{Y_l^m(\theta,\varphi)} \sqrt{\lambda(a,\theta,\varphi)}\psi_\Theta(\theta_{1/a},\varphi_{1/a}) d\Omega \\
&= \int_{\mathbb{S}^2} \overline{Y_l^m(\theta,\varphi)} \lambda^{-1/2}(a,\theta,\varphi)\psi_\Theta(\theta',\varphi') d\Omega'
\end{aligned} \tag{2.65}$$

To avoid confusion of the notation in (2.65) and also in later part, we clarify here that $\varphi(\theta',\varphi')$ are now taken as new variables while θ and φ shall be understood as $\theta(\theta',\varphi')$ and $\varphi(\theta',\varphi')$.

Insert (2.51) into (2.65) and replace $\sqrt{\lambda(a,\theta,\varphi)}$ by the expression $a^{-\frac{1+\alpha}{2}}(1+J^2)\cos^2(\frac{\theta'}{2})$ in (2.55), we get

$$(\widehat{(\psi_\Theta)_a})_{lm} = 2a^{\frac{1+\alpha}{2}} c_{l,m} \int_0^{2\pi}\int_0^{\pi} e^{-im\varphi} \sin^m(\theta) P_l^{(m)}(\cos\theta) \frac{\psi_\Theta(\theta',\varphi')\tan(\frac{\theta'}{2}) d\theta' d\varphi'}{1+J^2} \tag{2.66}$$

where $J = \tan(\frac{\theta'}{2})\cos\varphi' \sqrt{(a^2 + a^{2\alpha}\tan^2\varphi')}$.

Let us examine the integrand in (2.66) in detail. Since $\theta' \in (0,\Theta)$, $\tan\frac{\theta'}{2} \leqslant \tan\frac{\Theta}{2}$ bounded almost everywhere. What's more, Lemma 2.14 and the observation that $\theta \to 0$ as $a \to 0$ tell us $\sin^m \theta P_l^{(m)}(\cos\theta) = O(a^{\alpha m} l^{2m})$. Indeed, according to (2.52) there is

$$\begin{aligned}
|\sin\theta| &= \left|\frac{2\tan(\frac{\theta'}{2})\cos\varphi' \sqrt{a^2 + a^{2\alpha}\tan^2\varphi'}}{1+\tan^2(\frac{\theta'}{2})\cos^2\varphi'(a^2 + a^{2\alpha}\tan^2\varphi')}\right| \\
&= |2\tan(\frac{\theta'}{2})\cos\varphi' \sqrt{a^2 + a^{2\alpha}\tan^2\varphi'}| + O(a^{3\alpha}) \\
&= O(a^\alpha)
\end{aligned}$$

we conclude that for fixed l,

$$Y_l^m(\theta,\varphi) = O(a^{\alpha|m|+(1+\alpha)/2}) \tag{2.67}$$

for all $m \geqslant 0$, and hence for all $|m| \leq l$ due to the fact that

$$Y_l^{-m} = (-1)^m \overline{Y_l^m} \tag{2.68}$$

Thus $|(\widehat{(\psi_\Theta)_a})_{lm}| = O(a^{\alpha m+(1+\alpha)/2})$, $(\widehat{(\psi_\Theta)_a})_{lm} Y_l^m \in \mathcal{A}$ for all $|m| \geqslant 1$, and $\|\mathbb{P}_l(\psi_\Theta)_a\|_2^2 = \sum_{|m|\leq l} |(\widehat{(\psi_\Theta)_a})_{lm}|^2 = o(a^2)$ if and only if we impose the requirement that

$$\begin{aligned}
0 &= \lim_{a\to 0} a^{-1}(\widehat{(\psi_\Theta)_a})_{l0} \\
&= 2\lim_{a\to 0} a^{-1/4} c_{l,0} \int_0^{2\pi}\int_0^{\Theta} P_l(\cos\theta)\frac{\psi(\theta', \varphi')\tan(\frac{\theta'}{2})}{1+J^2} d\theta' d\varphi' \\
&= 2\lim_{a\to 0} a^{-1/4} c_{l,0} \int_0^{2\pi}\int_0^{\Theta} \psi(\theta', \varphi')\tan(\frac{\theta'}{2}) d\theta' d\varphi'
\end{aligned}$$

namely that $\int_0^{2\pi}\int_0^{\Theta}\psi(\theta', \varphi')\tan(\frac{\theta'}{2})d\theta' d\varphi' = 0$. The last step used Lebesgue dominated Theorem with the observation that $|P_l(\cos\theta)| \leqslant P_l(1) = 1$, $J \to 0$ as $a \to 0$. Although we used a specially designed ψ_Θ, the whole argument we used above applies to any function in $\mathcal{K}$, namely $\mathcal{K}\cap\mathcal{A} = \mathcal{K}\cap\mathcal{N}$.

Due to our assumption that $\psi \in \mathcal{B}_n$, $|\psi(\theta', \varphi')\tan^3(\frac{\theta'}{2})|$ is bounded by some $M > 0$ almost everywhere whenever $|\Theta - \pi| < \delta$, we see that if $\|\mathbb{P}_l(\psi)_a\|_2^2 = \lim_{\Theta\to\pi}\|\mathbb{P}_l(\psi_\Theta)_a\|_2^2 = o(a^2)$, we necessarily have

$$0 = \lim_{\Theta\to\pi}\int_0^{2\pi}\int_0^{\Theta}\psi(\theta', \varphi')\tan(\frac{\theta'}{2})d\theta' d\varphi' = \int_0^{2\pi}\int_0^{\pi}\psi(\theta', \varphi')\tan(\frac{\theta'}{2})d\theta' d\varphi'$$

Conversely, assume $\int_0^{2\pi}\int_0^{\pi}\psi(\theta,\varphi)\tan(\frac{\theta}{2})d\theta d\varphi = 0$. Let

$$\widetilde{\psi_\Theta}(\theta,\varphi) = \begin{cases} \psi_\Theta(\theta,\varphi) - \mathfrak{T}\psi_\Theta(\theta,\varphi) & \text{for } \theta \leq \Theta \\ 0 & \text{for } \theta > \Theta \end{cases} \tag{2.69}$$

where

$$\mathfrak{T}\psi_\Theta(\theta,\varphi) = (\tan\frac{\theta}{2})^{-1}\int_0^{2\pi}\int_0^{\pi}\psi_\Theta(\theta,\varphi)\tan(\frac{\theta}{2})d\theta d\varphi \tag{2.70}$$

Since $\widetilde{\psi_\Theta}(\theta,\varphi) \in \mathcal{K}_\Theta \cap \mathcal{N}$, by our former conclusion we get $\widetilde{\psi_\Theta}(\theta,\varphi) \in \mathcal{A}$. Hence it follows from Lemma 2.12 that $\|\mathbb{P}_l(\widetilde{\psi_\Theta})_a\|_2^2 = o(a^2)$. Fix $\theta < \pi$, it's easy to see that $\widetilde{\psi_\Theta}(\theta,\varphi) \to \psi(\theta,\varphi)$ as $\Theta \to \pi$, and that $|(\widetilde{\psi_\Theta}(\theta,\varphi) - \psi_\Theta(\theta,\varphi))\tan^3\frac{\theta}{2}|$

is bounded for any $\Theta < \pi$. In fact, there is

$$
\begin{aligned}
&|\widetilde{\psi_\Theta}(\theta,\varphi) - \psi(\theta,\varphi)| \tan^3(\frac{\theta}{2}) \\
&\leqslant \left(|\widetilde{\psi_\Theta}(\theta,\varphi) - \psi_\Theta(\theta,\varphi)| + |\psi_\Theta(\theta,\varphi) - \psi(\theta,\varphi)|\right) \tan^3(\frac{\theta}{2}) \\
&\leqslant \left|\int_0^{2\pi}\int_0^{\pi} (\psi_\Theta(\theta,\varphi) - \psi(\theta,\varphi)) \tan(\frac{\theta}{2}) d\theta d\varphi\right| \tan^2(\frac{\theta}{2}) \\
&+ \max_{\theta>\Theta} \left|\psi(\theta,\varphi)\tan^3(\frac{\theta}{2})\right| \\
&\leqslant \tan^2\left(\frac{\theta}{2}\right) \int_0^{2\pi}\int_\Theta^{\pi} \left|\psi(\theta,\varphi)\tan(\frac{\theta}{2})\right| d\theta d\varphi + \max_{\theta>\Theta}\left|\psi(\theta,\varphi)\tan^3(\frac{\theta}{2})\right|
\end{aligned}
$$

which is bounded almost everywhere as Θ approaches π, since $\psi \in \mathcal{B}$. By replacing $\widetilde{\psi_\Theta}$ in (2.66), we use Lebesgue dominated Theorem again to get

$$
\lim_{a\to 0} a^{-2}\|\mathbb{P}_l(\psi_a)\|_2^2 = \lim_{a\to 0}\lim_{\Theta\to\pi} a^{-2}\|\mathbb{P}_l(\widetilde{\psi_\Theta})_a\|_2^2 = 0
$$

In the last step the exchange of order of limits is allowable since the convergence is uniform. Indeed, for $m = 0$ and any $a \in (0, 1]$, there is

$$
a^{-1}\left|\widehat{\left((\widetilde{\psi_\Theta})_a\right)}_{l0} - \widehat{(\psi_a)}_{l0}\right| = a^{-\frac{1}{4}}\left|\int_0^{2\pi}\int_0^{\pi} \frac{P_l(\cos\theta)}{1+J^2}\left(\widetilde{\psi_\Theta} - \psi\right)(\theta',\varphi')\tan\frac{\theta'}{2} d\theta' d\varphi'\right|
$$

Since $\widetilde{\psi_\Theta} - \psi \in \mathcal{N}$, the latter goes to

$$
4a^{\frac{3}{4}}\left|\int_0^{2\pi}\int_0^{\pi} \frac{d}{da}\left(\frac{P_l(\cos\theta)}{1+J^2}\right)\left(\widetilde{\psi_\Theta} - \psi\right)(\theta',\varphi')\tan\frac{\theta'}{2} d\theta' d\varphi'\right|
$$

as $a \to 0$, which further converges to zero uniformly due to the fact that $|\widetilde{\psi_\Theta}(\theta,\varphi) - \psi(\theta,\varphi)|\tan^3\left(\frac{\theta}{2}\right)$ is bounded almost everywhere. For $m \neq 0$ it can be verified similarly. Thus we have proved that $\psi \in \mathcal{A}$.

□

This main result reduces the complicated admissibility condition into the easy-to-check condition in Definition 2.13, hence greatly helpful for selecting candidate shearlets. Another natural question is whether an admissible shearlet remains admissible after dilation. In other words, we need to check whether our system defined in (2.8) is closed under the dilation operation. The following lemma gives us a positive answer.

Proposition 2.16. *If $\psi \in \mathcal{N} \cap \mathcal{B}_n$, then $\psi_a \in \mathcal{N} \cap \mathcal{B}_n$ for any $a \in \mathbb{R}_+$ and $n \geqslant 1$*

Proof. Let $\theta^{'} = \theta_{1/a}$ and $\varphi^{'} = \varphi_{1/a}$. Suppose $|\psi(\theta,\varphi)\tan^n \frac{\theta}{2}| \leqslant M$ for some $M > 0$, then by (2.51) and Proposition 2.6,

$$\begin{aligned}|\psi_a(\theta,\varphi)\tan^n \frac{\theta}{2}| &= |\sqrt{\lambda(a,\theta,\varphi)}\psi(\theta^{'},\varphi^{'})\tan^n \frac{\theta}{2}| \\ &\leqslant \frac{2a^{\frac{3-\alpha}{2}+n}M}{|\Phi_a^- \cos\theta + \Phi_a^+|(\cos^2\varphi + a^{2-2\alpha}\sin^2\varphi)^{n/2}}\end{aligned}$$

Further, we observe that

$$\begin{aligned}&\int_0^{2\pi}\int_0^{\pi} \psi_a(\theta,\varphi)\tan(\frac{\theta}{2})d\theta d\varphi \\ &= \int_0^{2\pi}\int_0^{\pi} \sqrt{\lambda(a,\theta,\varphi)}\psi(\theta_{1/a},\varphi_{1/a})\tan(\frac{\theta}{2})d\theta d\varphi \\ &= \int_0^{2\pi}\int_0^{\pi} \frac{\sec^2\frac{\theta}{2}}{2}\lambda^{-1/2}(a,\theta,\varphi)\psi(\theta^{'},\varphi^{'})\sin\theta^{'}d\theta^{'}d\varphi^{'} \\ &= \int_0^{2\pi}\int_0^{\pi} \frac{1+J^2}{2}\frac{a^{(\alpha+1)/2}\sec^2(\frac{\theta^{'}}{2})}{1+J^2}\psi(\theta^{'},\varphi^{'})\sin\theta^{'}d\theta^{'}d\varphi^{'} \\ &= a^{(\alpha+1)/2}\int_0^{2\pi}\int_0^{\pi} \psi(\theta^{'},\varphi^{'})\tan(\frac{\theta^{'}}{2})d\theta^{'}d\varphi^{'}\end{aligned}$$

therefore $\psi_a(\theta,\varphi) \in \mathcal{B}_n \cap \mathcal{N}$.

□

Remark 2.17. Once we have a square integrable function ψ on the sphere which fulfills the requirements in Theorem 2.15, then taking the difference of two sides of (2.3), we see a natural candidate of admissible shearlets is the function $\psi_a(\theta,\varphi) - a^{3/4}\psi(\theta,\varphi)$. Besides, for hollow pole functions, operator $\mathcal{D}(a)$ preserves regularity on the whole sphere. Indeed, given $f \in C^k(\mathbb{S}^2)$ with $k \geqslant 0$, it is obvious that $\mathcal{D}(a)f$ is k-times continuously differentiable away from the pole according to (2.54); while if f is a hollow pole function, then its dilated version keeps regularity at $\theta = \pi$, hence $\mathcal{D}(a)f \in C^k(\mathbb{S}^2)$.

The relationship between functions on $\mathbb{R}^2$ and shearlets on $\mathbb{S}^2$ is described by the next proposition. More precisely, we prove that every zero mean function on $\mathbb{R}^2$ after being projected inversely by U is an admissible shearlet on the sphere.

Proposition 2.18. *Let $\psi \in L^2(\mathbb{R}^2)$ be a function such that*

$$\int_{\mathbb{R}^2} \psi(x_1, x_2) d\mathbf{x} = 0$$

and

$$|\psi(x_1, x_2) r^{n+2}| \leqslant M \quad \forall n \leqslant N$$

for some $M > 0$ and $N \in \mathbb{Z}^+$. Then $U^{-1}\psi \in \mathcal{B}_N \cap \mathcal{N} \in \mathcal{A}$.

Proof. Let $r = 2\tan(\frac{\theta}{2})$, ψ compactly supported implies $U^{-1}\psi$ is a hollow pole function.

$$\begin{aligned}
\int_0^{2\pi}\int_0^{\pi} U^{-1}\psi(\theta,\varphi)\tan(\frac{\theta}{2})d\theta d\varphi &= \int_0^{2\pi}\int_0^{\pi} \psi(r\cos\varphi, r\sin\varphi)\sec^2(\frac{\theta}{2})\tan(\frac{\theta}{2})d\theta d\varphi \\
&= \frac{1}{2}\int_0^{2\pi}\int_0^{\infty} \psi(r\cos\varphi, r\sin\varphi) r dr d\varphi \\
&= \int_{\mathbb{R}^2} \psi(x_1, x_2) d\mathbf{x}
\end{aligned}$$

Since $\int_{\mathbb{R}^2} \psi(x_1, x_2) d\mathbf{x} = 0$, by definition, we have $U^{-1}\psi \in \mathcal{N}$. Furthermore, $|\psi(x_1, x_2) r^{n+2}| \leqslant M$ for all $n \leqslant N$ implies that $U^{-1}\psi \in \mathcal{B}_N$. Therefore, $U^{-1}\psi$ is an element in $\mathcal{A}$. □

2.4 Other approaches: a selective review

In this section let us have a tour and make some comments on other representative and creative approaches of constructing spherical wavelets that exhibit certain merits as well as insufficiency in different applications.

Firstly, I want to comment that, in the dilation aspect, the anisotropic spherical Gaussian(ASG) is probably the most widely used tool by engineers, for instance in the description of directional dependence radio waves from antenna. Mathematically, without loss of generality, if we denote by $\mathbf{x}$,$\mathbf{y}$,$\mathbf{z}$ a set of mutually perpendicular unit eigenvectors of a 3×3 symmetric matrix A, corresponding to eigenvalues $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$ respectively, the ASG(as a function of ξ) is defined to be

$$G(\xi, A) = e^{\xi^T A \xi} \cdot max\{\mathbf{z} \cdot \xi, 0\}$$

Note that the matrix A can be rewritten in the form $A = (\lambda_1 - \lambda_3)\mathbf{x}\mathbf{x}^T + (\lambda_2 - \lambda_3)\mathbf{y}\mathbf{y}^T + \lambda_3 I$, hence ASG has another expression

$$e^{-[(\lambda_1-\lambda_3)(\xi\cdot\mathbf{x})^2+(\lambda_2-\lambda_3)(\xi\cdot\mathbf{y})^2+\lambda_3]} \cdot max\{\mathbf{z} \cdot \xi, 0\} \tag{2.71}$$

Comparing it with the traditional von Mises-Fisher distribution

$$\begin{aligned} e^{2\lambda(\xi\cdot\mathbf{z}-1)} &= e^{-\lambda[(\xi\cdot\mathbf{x})^2+(\xi\cdot\mathbf{y})^2]}\cdot e^{-\lambda(\xi\cdot\mathbf{z}-1)^2} \\ &= e^{-\lambda\left[(\xi\cdot\mathbf{x})^2+(\xi\cdot\mathbf{y})^2])+1\right]}\cdot e^{-\lambda(\xi\cdot\mathbf{z})^2+2\lambda(\xi\cdot\mathbf{z})} \end{aligned} \tag{2.72}$$

where λ is the bandwidth and $\mathbf{z}$ is chosen to be the lobe axis, the axis of the smallest eigenvalue which corresponds to the smallest radiation amplitude, we see that the last $\mathbf{z}$ term is replaced by a smoothing term, which is used to constrain value in upper hemisphere while preserving smoothness. It is demonstrated[7] that ASG have approximate closed-form solutions for product and convolution operators.

A subdivision scheme is used in the 1995 work by Schröder and Swelden[25] to build the orthogonal Haar wavelet transform on arbitrary manifolds. This method starts with an initial quasi-uniform triangulation of the surface and builds finer level mesh by connecting the midpoint of each edge. Whenever there is a multiresolution analysis on the surface either based on vertex basis or face basis, one can use the lifting scheme to obtain wavelets of better properties like improved smoothness or vanishing moments. That is to say, if we have a biorthogonal wavelet basis system such that for any $f \in L^2(\mathbb{S}^2)$ there is

$$\begin{aligned} f &= \sum_{j,m}\langle f, \tilde{\psi}_{j,m}\rangle \psi_{j,m} \\ \psi_{j+1,k} &= \sum_{k'}\tilde{h}_{j,k,k'}\varphi_{j,k'} + \sum_{m}\tilde{g}_{j,k,m}\psi_{j,m} \end{aligned} \tag{2.73}$$

with refinement relations $\sum\limits_{k'}\tilde{h}_{j,k,k'}\tilde{\varphi}_{j+1,k'} = \tilde{\varphi}_{j,k}$ and $\sum\limits_{m}\tilde{g}_{j,k,m}\tilde{\varphi}_{j+1,m} = \tilde{\psi}_{j,k}$, then a new wavelet system can be generated through the following design of filters

$$\begin{cases} h^{new}_{j,k,k'} = h_{j,k,k'} \\ \tilde{h}^{new}_{j,k,k'} = \tilde{h}_{j,k,k'} + \sum\limits_{m} c_{j,k,m}\tilde{g}_{j,k',m} \\ \tilde{g}^{new}_{j,k,m} = \tilde{g}_{j,k,m} \\ g^{new}_{j,k,m} = g_{j,k,m} - \sum\limits_{k'} c_{j,k,k'}h_{j,m,k'} \end{cases} \tag{2.74}$$

namely the scaling function remains the same while its dual and the wavelet function are lifted, where $g_{j,k,m}$ and $h_{j,m,k'}$ are similarly defined and $c_{j,k,k'}$ and $c_{j,k,m}$ are coefficients to be chosen.

There are in general two types of schemes of subdivision, the face splitting scheme and the vertex splitting scheme. In the face setting, the scaling function and its bi-orthogonal dual at different levels supported in a local triangle can be simply formed by taking the characteristic functions, namely setting $\varphi_{j,k} = \chi_{T_{j,k}}$ and $\tilde{\varphi}_{j,k} = \frac{1}{|T_{j,k}|}\chi_{T_{j,k}}$. If the subdivision triangles of $T_{j,k}$ are labeled as $\{T_0, T_1, T_2, T_3\}$ with T_0 indicating the middle triangle, the wavelet dual pairs can be chosen as

$$\psi_{j,m} = 2\left(\varphi_{j+1,m} - \frac{\int_{\mathbb{S}^2} \varphi_{j+1,m} d\Omega}{\int_{\mathbb{S}^2} \varphi_{j,0} d\Omega}\varphi_{j,0}\right) \tag{2.75a}$$

$$\tilde{\psi}_{j,m} = \frac{1}{2}(\tilde{\varphi}_{j+1,m} - \tilde{\varphi}_{j,k}) \tag{2.75b}$$

where $m = 1, 2, 3$, so that the wavelet here has vanishing integral. In order to achieve higher vanishing moments in the sense that there exist n linearly independent polynomials restricted on the sphere such that their wavelet coefficients vanish for all $j \geqslant 0$ and all m in index set at level j, one can propose the lifted dual as

$$\tilde{\psi}_{j,m} = \frac{1}{2}(\tilde{\varphi}_{j+1,m} - \tilde{\varphi}_{j,k}) - \sum_{k' \in \mathcal{N}(k)} c_{j,k',m}\tilde{\varphi}_{j,k'} \tag{2.76}$$

where $k^{'}$ are index of the neighboring triangles of $T_{j,k}$.

In the vertex setting, the scaling functions are often chosen as delta functions. If the vertices are labeled by k at level j and the midpoint of an edge(or the newly generated vertex) is labeled by m, then one could simply subsample followed by upsampling the scaling coefficients and let the refinement relation be

$$\psi_{j,m} = \varphi_{j+1,m} - \sum_{k \in \mathcal{N}(m)} c_{j,k,m}\varphi_{j,k}$$

A special choice would be

$$c_{j,k,m} = \begin{cases} \frac{\int_{\mathbb{S}^2} \varphi_{j+1,m} d\Omega}{2\int_{\mathbb{S}^2} \varphi_{j,k} d\Omega} & \text{for } k = v_1, v_2 \\ 0 & \text{otherwise} \end{cases}$$

where v_1 and v_2 are endpoints corresponding to m.

Suppose now we have a convex polyhedron Γ with triangular surfaces and having all its vertices located on the 2-sphere. Through certain subdivision scheme we obtain refined triangulations $\mathcal{T}^j$ with the set of vertices $\mathcal{V}^j$ and nodal functions ϕ_v at vertex $v \in \mathcal{V}^j$. The space $\mathbb{P}^1_j$ of piecewise linear continuous function on $\mathcal{T}^j$ is a subspace of $\mathbb{P}^1_{j+1}$. In fact, if only neighboring vertices are involved, there is the relation $\phi^j_v = \phi^{j+1}_v + \frac{1}{2}\sum_{v' \in \mathcal{N}(v)} \phi^{j+1}_{v'}$; while in the butterfly scheme more vertices could be involved. Thus the composition of ϕ_v^j and inverse projection $p^{-1} : \mathbb{S}^2 \to \partial\Gamma$, denoted by φ_v^j, satisfy the refinement equation

$$\varphi_v^j = \varphi_v^{j+1} + \frac{1}{2} \sum_{v' \in \mathcal{N}(v)} \varphi_{v'}^{j+1} \tag{2.77}$$

and if $u \in \mathcal{V}^{j+1} \backslash \mathcal{V}^j$ be the newly added midpoint on an edge $[v_1, v_2]$, the next step is to define a wavelet space such that $W_j \bigoplus \mathbb{P}_j^1 = \mathbb{P}_{j+1}^1$. One strategy used in [70] is to assume that the wavelet function at m is a linear combination of the nodal functions of the previous level in the neighborhood of v_1 and v_2, and vanishes $\mathbb{P}_j^1$. It is proved there if one define

$$\begin{aligned} \psi_{j,u} &= \left(c_1 \phi_{v_1}^{j+1} + \sum_{v' \in \mathcal{N}^{j+1}(v_1)} c_{v'} \phi_{v'}^{j+1} \right) + \left(c_2 \phi_{v_2}^{j+1} + \sum_{v' \in \mathcal{N}^{j+1}(v_2)} c_{v'} \phi_{v'}^{j+1} \right) \\ &= \psi_{j,u}^{(1)} + \psi_{j,u}^{(2)} \end{aligned} \tag{2.78}$$

by requiring that for any $v \in \mathcal{V}^j$

$$\langle \psi_{j,u}^{(1)}, \phi_v^j \rangle_{\partial\Gamma} = \begin{cases} (-1)2^{-2j}\gamma & v = v_1 \\ 2^{-2j}\gamma & v = v_2 \\ 0 & else \end{cases}$$

and

$$\langle \psi_{j,u}^{(2)}, \phi_v^j \rangle_{\partial\Gamma} = \begin{cases} (-1)2^{-2j}\gamma & v = v_2 \\ 2^{-2j}\gamma & v = v_1 \\ 0 & else \end{cases}$$

where γ is a given nonzero constant, so that $\langle \psi_{j,u}, \phi_v^j \rangle_{\partial\Gamma} = 0$, then $\psi_{j,u}^{(1)}$ and $\psi_{j,u}^{(2)}$ are uniquely determined, hence $\psi_{j,u}$.

Spherical Haar wavelets are later improved into a both orthogonal and symmetric basis in the work [46], where instead of using the geodesic bisectors they smartly designed the subdivision by employing a spherical trigonometry formula from a college and school book a century ago that is not well-known to the people nowadays, so that areas of the children triangles on the finer level are equal.

The authors of [82] project from the plane to the sphere the Mexican hat wavelet $\frac{1}{\sqrt{2\pi}}(2 - x^2)e^{-x^2/2}$, which is almost the Laplacian of a Gaussian, to form the so called spherical Mexican wavelet. Noticed by computer scientists that spherical Haar wavelets are constructed after subdivision and the refining scheme depending heavily on the connectivity of the mesh hence computationally expensive, the authors from [59] recently introduced a Mexican hat wavelet formulated in the frequency domain as an alternative choice, namely by defining the continuous wavelet on an a compact manifold as the derivative of the heat kernel $h_t(x, y)$ and the discrete version wavelet as its difference. It has the advantage being localized in both space and frequency domains. Precisely, let

$$\psi_t(x, y) = \sum_{k=0}^{\infty} \lambda_k e^{-\lambda_k t} \phi_k(x) \phi_k(y) \tag{2.79}$$

with $\{\phi_k\}$ are eigenfunctions of the Laplace-Baltrami operator on the manifold and λ_k the corresponding distinct eigenvalues, then the continuous wavelets transform of a square integrable function f on $\mathbb{M}$ is

$$\mathcal{W}_\psi f(x, t) = \int_M \psi_t(x, y) f(y) dy$$

and it has the inverse transform

$$f = \int_0^\infty \mathcal{W}_\psi f(x, t) dt + \hat{f}(0) \phi_0(x) \tag{2.80}$$

By definition heat kernel and the associated wavelets have the properties that for all x and y on $\mathbb{M}$

$$\begin{cases} \lim_{t \to 0} h_t(x, y) = \delta_x(y) \\ \lim_{t \to \infty} h_t(x, y) = 1/\nu(M) \\ \lim_{t \to \infty} \psi_t(x, y) = 0 \\ \lim_{t \to 0^+} \psi_t(x, y) = \sum_k \lambda_k \phi_k(x) \phi_k(y) \end{cases} \tag{2.81}$$

In particular the last term is the kernel of $\Delta_{\mathbb{M}}$.

Similarly, if we divide the time line into a sequence $[t_0, \cdots, t_N]$, the discrete Mexican hat wavelet transformation and its inverse are

$$\begin{aligned} \widehat{\mathcal{W}_\psi^\dagger f}(k) &= \widehat{\psi_{t_j}}(k) \widehat{f}(k) \\ f(x) &= \sum_{j=1}^{N} \mathcal{W}_\psi^\dagger f(x, t_j) + \int_M h_{t_N}(x, y) f(y) dy \end{aligned} \tag{2.82}$$

where $\widehat{\psi_{t_j}}(k) = e^{-\lambda_k t_{j-1}} - e^{-\lambda_k t_j}$, using the fact that h_0 is the reproducing kernel of the Hilbert space $L^2(\mathbb{M})$. Since the heat kernel preserves the integral

$\int_{\mathbb{M}} h_t(x,y)dy = 1$ for all $t \geqslant 0$, the wavelet formulated in this way automatically has zero mean. I would like to briefly mention that a comparable construction of the Mexican hat wavelet is the Poisson wavelet from [51], where frame properties are discussed in a more theoretical way.

Another relatively new method is the Geometric Multi-Resolution Analysis(GMRA) adapted to a m-dimensional smooth compact Riemannian manifold $\mathbb{M}$ embedded in $\mathbb{R}^n$ was built in [97]. This method consists of the following steps. Firstly decompose the m-dimensional manifold in each scale $j \in \mathbb{Z}$ into subsets $U_j = \{U_{j,k}\}_{k \in \mathcal{K}_j}$, such that:

(i) They completely cover $\mathbb{M}$ and $\nu(U_{j,k'} \bigcap U_{j,k}) = 0$ for any $k' \neq k \in \mathcal{K}_j$

(ii) $U_{j',k'}$ is included in one and only one $U_{j,k}$ whenever $j < j'$

(iii) comparison principle: for each level j,

$$B_\rho(c_{j,k}, 2^{-j'}) \subset U_j \subset B_\rho(c_{j,k}, 2^{-j})$$

holds for some $j' > j$, where the Riemannian metric is denoted by ρ and the Borel measure by μ. Such a decomposition for instance can be achieved through intersection of the manifold with dyadic cubes in $\mathbb{R}^n$.

The second step is to find the minimizer of the functional

$$\min_P \int_{U_{j,k}} d(x,P)^2 d\nu(x) \tag{2.83}$$

where d measures the Euclidean distance to an affine plane P. The solution of this minimization problem is the affine space spanned by the m eigenvectors corresponding to the maximum eigenvalues of the covariance matrix

$$\mathbb{E}[(x - \mathbb{E}x)(x - \mathbb{E}x)^\dagger | x \in U_{j,k}] = B_{j,k}\Sigma_{j,k}B_{j,k}^\dagger \tag{2.84}$$

with Σ_{jk} a $\dim P \times \dim P$ diagonal matrix and $\dim P$ equals the number of vectors in $B_{j,k}$, centered at

$$c_{j,k} = \mathbb{E}[x \in U_{j,k}] = \frac{1}{\nu(U_{j,k})} \int_{U_{j,k}} x d\nu(x) \in \mathbb{R}^n$$

denoted by $c_{j,k} + V_{j,k}$. Indeed, let ξ_i $(i = 1, \cdots, n-m)$ be orthonormal unit vectors in $\mathbb{R}^n$, then (2.83) equals to

$$\begin{aligned}
&\min_{\substack{\lambda\in\mathbb{R}^{n-m}\\ \xi_i\in\mathbb{S}^{n-1}}} \frac{1}{\nu(U_{j,k})}\int_{U_{j,k}}\sum_{i=1}^{n-m}|x\cdot\xi_i-\lambda_i|^2 d\nu(x)\\
&=\min_{\xi_i\in\mathbb{S}^{n-1}} \frac{1}{\nu(U_{j,k})}\int_{U_{j,k}}\sum_{i=1}^{n-m}|\langle x-\mathbb{E}x,\xi_i\rangle|^2 d\nu(x)\\
&=\min_{\xi_i\in\mathbb{S}^{n-1}}\sum_{i=1}^{n-m}\xi_i^\dagger\mathbb{E}[(x-\mathbb{E}x)(x-\mathbb{E}x)^\dagger|x\in U_{j,k}]\xi_i
\end{aligned} \tag{2.85}$$

therefore the minimum value is achieved by choosing arbitrary $n-m$ eigenvectors corresponding to the minimum eigenvalues. We point out that in the paper [97] the assumption that $\mathbb{E}[(x-\mathbb{E}x)(x-\mathbb{E}x)^\dagger|x\in U_{j,k}]$ has rank $m\ll n$ would be not right when the manifold is of poor regularity or having large curvature. Covariance matrix in general could have full rank when its curvature is not identically zero. In fact, assume without loss of generality that $B_{j,k}$ form the last m coordinates and points in U have the expression $x=(h_1(x^{'}),\cdots,h_{n-m}(x^{'}),x^{'})$, where $x^{'}$ is m-tuple $(x_{n-m+1},\cdots,x_n)$. If at U the curvature is large, then $U\subset P$ and for any $i\leqslant n$ and $j>m$, since $x_j=0$ there is

$$c_{ij}=\frac{1}{|U|}\int_U(x_j-\mathbb{E}x_j)(x_i-\mathbb{E}x_i)dx=0$$

hence the rank of $\mathbb{E}[(x-\mathbb{E}x)(x-\mathbb{E}x)^\dagger|x\in U]$ is at most m. However, when the curvature at U is not zero, say $h_i>0$ for all $i\leqslant n$, then $c_{ii}>0$ and $\det\mathbb{E}[(x-\mathbb{E}x)(x-\mathbb{E}x)^\dagger|x\in U]$ does not have to vanish. In other words covariance matrix can have full rank n regardless of the fact that $\mathbb{M}$ is a m-dim manifold. Nevertheless if we know in advance that the surface consists of points distributed around a hyperplane or its sectional curvatures are small everywhere, namely $h_i\ll\operatorname{diam}(U)$, then the low rank assumption holds as a first order approximation.

Let $\Pi_{j,k}:=B_{j,k}B_{j,k}^\dagger$, the third step is to compare the difference between the affine projection $x_{j,k}=\Pi_{j,k}(x-c_{j,k})+c_{j,k}$ to the plane centered at $c_{j,k}$ and the projection $\Pi_{j+1,k'}(x-c_{j+1,k'})+c_{j+1,k'}$ to the next level plane, where

$x \in U_{j,k} \cap U_{j+1,k'}$. Without confusion we omit the index k and $k^{'}$. Then

$$\begin{aligned}
\Delta_j x &:= \Pi_{j+1}(x - c_{j+1}) - \Pi_j(x - c_j) \\
&= (I - \Pi_j)\Pi_{j+1}(x - c_{j+1}) - \Pi_j[(x - c_j) - \Pi_{j+1}(x - c_{j+1})] \\
&= (I - \Pi_j)\Pi_{j+1}(x - c_{j+1}) - \Pi_j[(x - c_j) - \Pi_{J+1}(x - c_{J+1})] \\
&+ \sum_{s=j+1}^{J} \Pi_j(\Pi_{s+1}(x - c_{s+1}) - \Pi_s(x - c_s)) \\
&= w_j - \Pi_j[(x - c_j) - \Pi_{J+1}(x - c_{J+1})] + \Pi_j \sum_{s=j+1}^{J} \Delta_s x
\end{aligned} \tag{2.86}$$

for any $J \geqslant j + 1$, where $w_j \in W_{j+1} := (I - \Pi_j)V_{j+1}$. Thus we have the decomposition of $\Pi_{j+1}(x - c_{j+1}) \in V_{j+1}$ into two orthogonal parts, namely the detail part w_j in W_j and the coarse part that belongs to V_j. An obvious fact in this method is that the centers change at different levels and eigenvectors must be recalculated, as a result there is no simple operations like translation and dilation directly applicable.

In my humble opinion one of the main obstacles in local projection approach is how to smartly build an adaptive mesh on an arbitrary surface so that the transition functions have good regularity properties. Even for the special case of the unit sphere, it is not a trivial question. Best performance in regularity aspect is achieved interestingly, however, by the earliest works in Germany. In [74] the authors divide the whole domain into quad mesh and utilize tensor product of exponential splines and B-wavelets to form the wavelets on a square and join smoothly the pull-back of those wavelets which are able to obtain C^1 regularity at the pole and C^∞ elsewhere. In [79] arbitrary smoothness of the wavelets is realized on the whole sphere in theory, although it had the drawback in implementation according to Prof. Dahmen and Prof. Schneider. An optimized and implementable version is given by Kunoth and Sahner[48]. A similar strategy using local plane approximation for arbitrary manifold is adopted recently in [94]. The representation system formed in this way in general could be highly redundant and it is not always clear whether frame properties hold globally.

In our projection means for constructing spherical α-wavelets/shearlets, however, as I have mentioned in Remark 2.17, hollow pole functions preserve regularity globally under dilation and as you shall see in the next chapter, they generate a frame system under suitable assumptions, so that they recover a function completely and stably even though there might exist overlaps.

Those methods mentioned above as well as ours have potentially many collaborators in different scientific fields from astrophysics to medical imaging. For instance, as a strong support for the big bang theory, the cosmic microwave background was originally assumed to be perfectly uniform or isotropic[32]. However, detection of small fluctuation of its temperature leads physicists to search for the anisotropic structure behind it. Thus precise measurement of small parameters of the cosmological model is crucial according to physicists, and the scale property of wavelets naturally perfectly fits into this needs. In this aspect I would like to refer the readers to [24][58] for physics background, [42] for real observation data based on wavelet tools and [23] for a recent survey. In [90] spherical wavelets are used for analyzing the lithosphere structure of terrestrial planets including the Earth, Venus, Mars and Moon, where the admittance and correlation functions of given wavelet degree possess negative values for lowland basins and positive values for highlands. Besides, spherical wavelets are also used in image segmentation[84], and other applications.

Chapter 3

Extension to Miscellaneous Results

3.1 Kernel approximation

A real continuous kernel $K : \mathbb{M} \times \mathbb{M} \to \mathbb{R}$ for some subset $\mathbb{M} \subset \mathbb{R}^n$ is said to be *positive definite* if for any integer N the quadratic form

$$\sum_{i,j=0}^{N} K(x_i, x_j) c_i c_j \geqslant 0 \tag{3.1}$$

holds for arbitrary set of points $\{x_i\}_{i=1}^N$ in the set $\mathbb{M}$ and coefficient vector $\mathbf{c} = (c_1, c_2, \ldots, c_N) \in \mathbb{R}^N$. When K is symmetric, according to Theorem 4.7, integral operator $\int_{\mathbb{M}} K(x,y) f(y) dy$ has eigenfunctions $\{\varphi_n\}_{n\in\mathbb{N}}$ forming an orthonormal basis of $L^2(\mathbb{M})$, hence $K(x,y) = \sum_n c_n(x)\varphi(y)$ for some vector of functions $\mathbf{c}$ depending on x. Replacing this into the equation

$$\int_{\mathbb{M}} K(x,y)\varphi_n(y)dy = \lambda_n \varphi_n(x)$$

and using the linear independence of $\{\varphi_n\}_{n\in\mathbb{N}}$, we see that $c_n(x) = \lambda_n \varphi_n(x)$ and hence obtain the representation

$$K(x,y) = \sum_{n=0}^{\infty} \lambda_n \varphi_n(x) \varphi_n(y) \tag{3.2}$$

Conversely if K has such an expression with respect to an orthonormal basis, then each basis element φ_n is an eigenfunction corresponding to λ_n. Especially when $\mathbb{M}$ is equipped with a norm and $K(x,y) = \widetilde{K}(\|x-y\|)$ for some continuous function $\widetilde{K}$ and norm $\|\cdot\|$, it is called a *radial kernel*.

On a compact domain, Mercers' Theorem says that a continuous symmetric kernel K is a positive definite if and only if

$$\int_{\mathbb{M}\times\mathbb{M}} K(x,y)w(x)w(y)dxdy \geqslant 0 \tag{3.3}$$

for all $w \in L^1(\mathbb{M})$.

We remark here that both the kernel and the coefficient vector $\mathbf{c} \in \mathbb{R}^n$ can be complex, and the definition can obviously be rewritten in the matrix form $\mathbf{c}^\dagger K\mathbf{c} \geq 0$. Nevertheless the definition (3.1) shall serve our current purpose.

Let us list some important examples of positive definite kernels:

(i)(*Stationary kernel and its multivariate version*)

The name "stationary" comes from the translation-invariant property. Let $x, y \in \mathbb{R}^n$, a stationary kernel is of the form $K(x,y) = \tilde{k}(x-y)$ for some continuous function $\tilde{k} : \mathbb{R}^n \to \mathbb{R}$. Bochner showed that

$$\int_{\mathbb{R}^n} \tilde{k}(x-y)w(x)w(y)dxdy \geqslant 0$$

for all $w \in L^1(\mathbb{R})$ is equivalent to the existence of a positive, finite Borel measure μ such that $\tilde{k} = \hat{\mu}$.

In the multi-dimension situation tensor product of one dimension kernels $\tilde{k}_i : \mathbb{R} \to \mathbb{R}$ with $i = 1, \cdots, n$ leads to the multivariate kernel $K(x,y) = \prod_{i=1}^{n} \tilde{k}_i(x_i - y_i)$. This kind of kernels, since it involves different dimensions, could certainly embrace anisotropic traits, and is mostly used in statistics as a special nonparametric regression method. In that context, given d-variate random vectors $\mathbf{x}_i$ $(i = 1, \cdots, n)$ with a common density function ρ, one needs to wisely choose symmetric and positive definite $d \times d$ smoothing matrix H, such that the kernel density estimate

$$\rho_H(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} K_H(\mathbf{x} - \mathbf{x}_i)$$

with $K_H(\mathbf{x}) = |H|^{-1/2}K(H^{-1/2}\mathbf{x})$ minimizing the mean integrated squared error $\mathrm{MISE}(\rho_H) = \int E\left[(\rho_H(\mathbf{x}) - \rho(\mathbf{x}))^2\right] d\mathbf{x}$. The latter can be decomposed and written as

$$\begin{aligned}
&\int \left[\mathrm{Var}(\rho_H(\mathbf{x})) + (E\rho_H(\mathbf{x}) - \rho(\mathbf{x}))^2\right] d\mathbf{x} \\
&= \int \left[\frac{1}{n}\left(K_H^2 * \rho(\mathbf{x}) - E^2(\rho_H(\mathbf{x}))\right) + (K_H * \rho(\mathbf{x}) - \rho(\mathbf{x}))^2\right] d\mathbf{x}
\end{aligned}$$

where we used the fact that

$$\begin{aligned}
E\rho_H(\mathbf{x}) &= \int K_H(\mathbf{x}-\mathbf{y})\rho(\mathbf{y})d\mathbf{y} \\
&= \int K_H(\mathbf{y})[\rho(\mathbf{x}) - H^{-1/2}\mathbf{y}\cdot\nabla\rho(\mathbf{x}) + \frac{1}{2}(H^{1/2}\mathbf{y})^T \mathrm{Hess}.\rho(\mathbf{x}) H^{1/2}\mathbf{y} \\
&+ o((H^{1/2}\mathbf{y})^T(H^{1/2}\mathbf{y}))]d\mathbf{y}.
\end{aligned}$$

and

$$\begin{aligned}
\mathrm{Var}(\rho_H(\mathbf{x})) &= \frac{1}{n^2}E\left[\left(\sum_{i=1}^n K_H(\mathbf{x}-\mathbf{x}_i)\right)^2\right] - E\rho_H(\mathbf{x})E\rho_H(\mathbf{x}) \\
&= \frac{1}{n}\left[K_H^2 * \rho(\mathbf{x}) - (E\rho_H(\mathbf{x}))^2\right]
\end{aligned}$$

An example is that H chosen to be $\mathrm{diag}(h_1^2, \cdots, h_d^2)$, K_H to be the Gaussian $|H|^{-1/2}\frac{1}{(2\pi)^{d/2}}\exp(-\frac{1}{2}\mathbf{x}^T H^{-1}\mathbf{x})$, and

$$\rho_H(\mathbf{x}) = \frac{1}{n}\prod_j h_j^{-1}\sum_{i=1}^n K(\frac{x_1 - X_i^1}{h_1}, \cdots, \frac{x_d - X_i^d}{h_d})$$

Under the moments assumption that

$$\int K(\mathbf{x})d\mathbf{x} = 1, \int \mathbf{x}K(\mathbf{x})d\mathbf{x} = 0 \text{ and } \int K(\mathbf{x})\mathbf{x}\mathbf{x}^T d\mathbf{x} = \mu I$$

where $\mu = \int x_i^2 K(\mathbf{x})d\mathbf{x}$ independent of i, and the restriction that entries of H and $n^{-1}|H|^{-1/2}$ both go to zero as n goes to infinity, as well as that each term of Hess.ρ is piecewise continuous and square integrable, the above expressions become

$$\begin{aligned}
E\rho_H(\mathbf{x}) &= \rho(\mathbf{x}) + \frac{\mu}{2}tr(H.\mathrm{Hess}.\rho) + o(|H|) \\
\mathrm{Var}\rho_H(\mathbf{x}) &= \frac{1}{n}|H|^{-1/2}\int K(\mathbf{y})^2 d\mathbf{y}\rho(\mathbf{x}) + o(|H|^{-1/2}\cdot\frac{1}{n})
\end{aligned}$$

thus

$$\mathrm{MISE}(\rho_H) = \mathrm{AMISE}(\rho_H) + o(\frac{1}{n}|H|^{-1/2} + \mathrm{tr}^2 H)$$

where the leading term

$$\mathrm{AMISE}(\rho_H) = \frac{1}{n}|H|^{-1/2}\int K^2(\mathbf{x})d\mathbf{x} + \frac{\mu^2}{4}\int \mathrm{tr}^2(H\mathrm{Hess}.\rho)d\mathbf{x}$$

is called asymptotic mean integrated squared estimate. Certainly one can adopt an adaptive strategy and locate different smoothing matrices at different points, but in general a linear transformation enables one to consider random data of zero mean and unit covariant matrix. To maintain the structure of this chapter I do not intend to stretch out to full details, but rather refer the interested readers to monographs [15][72] among many excellent literatures. However, I do want to mention that the spirit of this method is very close to that of wavelets approximation, and it has the advantage that data distribution can be expressed graphically and perceived by human cognition very well, similar to the isotherm on a temperature distribution map.

(ii)(*Power series kernel and zonal kernel*)
For x,$y \in (-1,1)^n$, the kernel

$$K(x,y) = \sum_{\alpha\in\mathbb{N}_0^n} c_\alpha \frac{x^\alpha}{\alpha!}\frac{y^\alpha}{\alpha!} \quad \text{with} \quad \sum_{\alpha\in\mathbb{N}_0^n} \frac{c_\alpha}{(\alpha!)^2} < \infty \tag{3.4}$$

is firstly introduced in [1] as a generalization of the infinite product kernel $\sum\limits_{n=0}^{\infty} c_n(x\cdot y)^n$ given in [45], where the coefficients c_n are often assumed to be positive to guarantee the positive definiteness of the kernel, and the nonlinearly factorizable kernel of the form $\prod\limits_{j=1}^{d}\sum\limits_{n=0}^{\infty} c_n(x_j\cdot y_j)^n$. Among the simplest and most popular examples there are the exponential kernel $exp(x\cdot y) = \sum\limits_{n=0}^{\infty}\frac{x\cdot y}{n!} = \sum\limits_{\alpha\in\mathbb{Z}^n}\frac{1}{|\alpha|!}x^\alpha y^\alpha$ and $\prod\limits_{j=1}^{d}\frac{1}{1-cx_j\cdot y_j}$ with $c\in(0,1)$. It has the expansion (3.2) with the eigenfunctions being some modification of the Hermite polynomials H_j(see Example 4.13).

Zonal kernel adapts the stationary kernel in Euclidean spaces to spheres, in the sense that it uses the geodesic distance between x and y on $\mathbb{S}^{n-1}$ instead of the Euclidean distance. Note that when the kernel function is analytic, it reduces to an infinite product kernel. A special example is the spherical Gaussian

$exp(-2\epsilon(1-x\cdot y))$. We shall discuss properties of this kind of kernels further in this chapter.

(iii)(*Multiscale kernel*)

Sometimes the function spaces arising from PDEs or analysis have multiscale structure, hence a class of kernels combined with multiscale property is of particular interest. As was introduced in [34][35], the multiscale kernel on the plane has the form

$$K(x,y)=\sum_{j\geqslant 0}\lambda_j\sum_{k\in\mathbb{Z}^d}\varphi(2^jx-k)\varphi(2^jy-k) \quad \text{with} \quad x,y\in\mathbb{R}^d \tag{3.5}$$

It is clearly a special example of (3.2) and being particularly interesting in the sense that it is endowed with wavelet-like properties, where φ is either compactly supported or of decay rate $O(1+\|x\|^{\frac{-(d+1)}{2}})$. However, in the light of radial basis construction of spherical wavelets in section 2.1, I would like to propose a new type of kernels of the form

$$K_{\Phi,\lambda}(x,y)=\sum_{l\geqslant 0}\sum_{j\geqslant 0}\frac{1}{\mu_j}\widehat{\Phi}_j^4(l)P_l(x\cdot\xi_0)P_l(y\cdot\xi_0) \tag{3.6}$$

which can be viewed as spherical version of the multiscale kernel. As we shall see soon this kernel is a reproducing kernel for a certain Hilbert space.

(iv)(*Reproducing kernel*)

Suppose $\mathscr{H}$ is a function space which is Hilbert under certain norm, for instance a subspace of $L^2(\mathbb{M},\mathbb{R})$ or $C(\mathbb{M},\mathbb{R})$. It is well known that $\mathscr{H}$ has a reproducing kernel K if and only if the evaluation operator is bounded, namely there exists $C>0$ such that $|v(x)|\leqslant C\|v\|_{\mathscr{H}}$ for all x and any $v\in\mathscr{H}$. Reproducing kernel is positive definite due to the observation that

$$\sum_{i,j=1}^{N}c_ic_jK(x_i,x_j)=\sum_{i=1}^{N}\|c_iK_{x_i}\|^2\geqslant 0 \tag{3.7}$$

where we adopt the notation $K_x=K(x,\cdot)$. In particular, under the further assumption that

$$\sum_{i=1}^{N}c_iv(x_i)=0 \text{ for every } v \text{ and } \{x_i\}_i^N\text{implies } c_i=0 \tag{3.8}$$

K is strictly positive definite. Conversely, if K is symmetric positive definite on a set $\mathbb{M}$, there exists a unique Hilbert space in which K is the reproducing kernel. In fact, $\mathscr{H}$ can be defined as the completion of $\{K_x\}_{x\in\mathbb{M}}$ with respect to inner product $\langle \sum\limits_i a_i K_{x_i}, \sum\limits_j b_j K_{x_j}\rangle = \sum\limits_i \sum\limits_j a_i b_j K(x_i, x_j)$ and uniqueness of $\mathscr{H}$ follows from a mutual inclusion argument.

Many questions can be solved or approximated once they are restricted to a finite dimensional subspace which has a reproducing kernel. In applied mathematics, a useful example is that the minimization problem

$$\min_{g\in X}\{\|Ag - f\|_Y^2 + \alpha M(g)\} \tag{3.9}$$

in some Hilbert space X provided that M is a differentiable functional, can be solved explicitly by taking derivative if the solution is of the form $\sum\limits_{i=1}^{N} c_i K_{x_i}$. This kind of minimization is widely applicable in inverse problems[37]. In complex analysis, an important example is the Bermann kernel $K_B(z,\bar{w}) = \sum\limits_{k=0}^{\infty} \frac{1}{\lambda_k}\varphi_k(z)\overline{\varphi_k(w)}$ with $\{\varphi_k\}_{k\geqslant 0}$ an orthonormal basis such that

$$\int_{\mathfrak{C}} \varphi_k(z)\overline{\varphi_{k'}(z)}dz = \delta_{kk'} \quad \text{and} \quad \int_{\Omega} \varphi_k(z)\overline{\varphi_{k'}(z)}dz = \lambda_k\delta_{kk'}$$

which is the reproducing kernel for Bergmann space which consists of analytic functions on a bounded domain $\Omega \subset \mathbb{C}^d$ with finite L^2 norm; while the Hua-Poisson kernel $H(z,w) = \frac{P(z,\bar{w})P(w,\bar{z})}{P(z,\bar{z})}$ defined by Hua[57] with $P(z,\bar{w}) = \sum\limits_{n=0}^{\infty} \varphi_n(z)\overline{\varphi_n(w)}$, is the reproducing kernel for the class of functions $v(z) = \sum\limits_{k\geqslant 0} c_k\varphi_k(z)/P(z,\bar{z})$ on the characteristic manifold $\mathfrak{C} \subset \partial\Omega$ of Ω, namely

$$v(z) = \int_{\mathfrak{C}} H(z,w)v(w)dw \tag{3.10}$$

In particular, for Laplacian $\Delta_{B_{\mathbb{C}}} = 4\sum\limits_{j,k} \frac{1-|z|^2}{3}(\delta_{j,k} - \bar{z}_j z_k)\frac{\partial^2}{\partial z_k \partial \bar{z}_j}$ on the complex ball $B_{\mathbb{C}} \subset \mathbb{C}^n$ with $\mathfrak{C} = \mathbb{S}_{\mathbb{C}}^{d-1}$,

$$u(z) = \int_{\mathbb{S}_{\mathbb{C}}^{d-1}} H(z,w)f(w)dw \tag{3.11}$$

solves the Cauchy-Dirichlet problem

$$\begin{cases} \Delta_B u(z) = 0 & z \in B_{\mathbb{C}} \\ u|_{\mathbb{S}_{\mathbb{C}}^{d-1}}(z) = f(z) & z \in \mathbb{S}_{\mathbb{C}}^{d-1} \end{cases} \tag{3.12}$$

with $P_{B_{\mathbb{C}}}(z,w)=\frac{(d-1)!}{2\pi^d(1-z\cdot\bar{w})^d}$. Nevertheless, we shall not extend further into those examples.

In spherical context, note that $\mathcal{H}_l$ are RKHS themselves with kernel Z_l in (1.29). Suppose K is a reproducing kernel of a subspace $\mathscr{H}^{'}$ of the square integrable functions on the sphere and every Z_l is an element of $\mathscr{H}^{'}$, then K meets the condition (3.8), hence is strictly positive definite. Indeed, if we assume that $\sum\limits_{i=1}^{N} c_i v(x_i)=0$ for every $v\in\mathscr{H}^{'}$, in particular $\sum\limits_{i=1}^{N} c_i P_l(x_i\cdot x)=0$, then there is

$$\sum_{i=1}^{N} c_i Q_r(x_i\cdot x)=\sum_{l=0}^{\infty}\frac{2l+1}{4\pi}r^l\sum_{i=1}^{N} c_i P_l(x_i\cdot x)=0$$

where $Q_r(t)$ is the Poisson kernel. Let

$$\theta_i(x)=\begin{cases}\frac{x\cdot x_i-1+h}{h} & \text{if } x\cdot x_i<1-h\\ 0 & \text{else}\end{cases}\tag{3.13}$$

with $h\leqslant\min\limits_{i\neq j} x_i\cdot x_j$. Since θ_i is continuous and $\theta_i(x_j)=\delta_{ij}$, by Lemma 1.2 it follows that

$$c_i=\sum_{j=1}^{N} c_j\theta_{x_i}(x_j)=\sum_{j=1}^{N} c_j\lim_{r\to 1^-}\langle Q_r(x_j\cdot x),\theta_{x_i}(x)\rangle_{L^2}=0$$

A function $\psi:[0,\pi]\to\mathbb{R}$ is called *conditionally strictly positive definite* of order m on $\mathbb{S}^{n-1}$ if $\psi(x_i\cdot x_j)$ is positive definite with respect to

$$\{(c_1,\ldots,c_N):\sum_{i=1}^{N} c_i Y(x_i)=0\ \text{ for all } Y\in\mathcal{H}_{l-1}\}$$

a concept facilitates the discussion of positiveness on subspaces like $\bigoplus\limits_{l'\in L'}\mathcal{H}_{l'}$. Clearly a radial function ψ that is conditionally strictly positive definite on $\mathbb{R}^n$ is conditionally strictly positive definite on $\mathbb{S}^{n-1}$. Furthermore, $\sum\limits_{l'\in L'} a_{l'}P_{l'}(x_i\cdot y_j)$ being strictly positive definite for $a_{l'}>0$ is obviously equivalent to that of $\sum\limits_{l'\in L'} P_{l'}(x_i\cdot y_j)$, which holds if and only if $L^{'}$ contains infinitely many odd terms and even terms, which was proved in [6].

Before we proceed, I would like to give the following lemma which bridges two different reproducing kernel spaces. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be two reproducing kernel Hilbert spaces of real(or complex)-valued functions on $\mathbb{M}_1$ and $\mathbb{M}_2$ respectively, and K_1 be the reproducing kernel of $\mathscr{H}_1$. The lemma says that under certain isometries K_2 is uniquely determined by K_1.

Lemma 3.1. *If K_1 is the reproducing kernel of $\mathscr{H}_1$ and $F : \mathscr{H}_1 \to \mathscr{H}_2$ is an isometry such that $F\mathcal{I}_v h = \mathcal{I}_v F h$ for any integrable function $h : \mathbb{M}_1 \times \mathbb{M}_1 \to \mathbb{R}$ where $\mathcal{I}_v(\cdot) := \langle v, \cdot\rangle$ is the evaluation operator at $v \in \mathscr{H}_1$, then $FF_yK_1(x)$ is the reproducing kernel of $\mathscr{H}_2$.*

Proof. For fixed point y in the domain $\mathbb{M}_2$ of the functions in $\mathscr{H}_2$, $(F_yK_1)(z) := F(K_1(\cdot, z))(y)$ is a function in $\mathscr{H}_1$ with respect to $x \in \mathbb{M}_1$.

$$\begin{aligned}\langle Fv, F(F_yK_1)\rangle &= \langle v, F_yK_1\rangle \\ &= \mathcal{I}_v(F_yK_1) \\ &= (F\mathcal{I}_vK_1)(y) \\ &= Fv(y)\end{aligned} \tag{3.14}$$

By uniqueness of the reproducing kernel we arrive at our conclusion. □

Remark 3.2. Whenever there is a linear isometry F between $L^2(\mathbb{S}^2)$ into $L^2(\mathbb{M})$, we are able to use the the above lemma to obtain the reproducing kernel of $L^2(\mathbb{M})$. In fact, the linearity of F implies that F commutes with $\mathcal{I}_v$, the integral operator.

For sampling points $x_1, \cdots, x_N$ in $\mathbb{M}$, it is a classical result that v^* from a RKHS $\mathscr{H}$ achieves $\sup\limits_{\|v\|^2 \leqslant E} \sum\limits_{i=1}^{N} v^2(x_i)$ when

$$v^*(x) = \sum_{i=1}^{N} \xi_M(x_i) K_{x_i} \tag{3.15}$$

with ξ_M the eigenfunction corresponding to the maximum eigenvalue λ_M of the matrix $(K(x_i, x_j))_{i,j=1,\cdots,N}$. In fact,

$$\begin{aligned}\sum_{i=1}^{N} v^2(x_i) &= \sum_{i=1}^{N} v(x_i)\langle v, K_{x_i}\rangle \\ &\leqslant \|v\| \sqrt{\sum_{i,j=1}^{N} v(x_i)v(x_j)K(x_i, x_j)} \\ &\leqslant E\sqrt{\lambda_M \sum_{i=1}^{N} \xi_M^2(x_i)}\end{aligned} \tag{3.16}$$

with equality holds if and only if $v = c\sum_{i=1}^{N} v(x_i)K_{x_i}$ and $\sum_{j=1}^{N} v(x_j)K(x_i, x_j) = \lambda_M v(x_i)$, namely $v = v^*$.

Another question that is often asked is whether there exist $v \in \mathscr{H}$ of minimum norm that solves the equations

$$v(x_i) = a_i \tag{3.17}$$

where $a_i \in \mathbb{R}$ are given. The answer is affirmative at the presence of a reproducing kernel that satisfies (3.8). Indeed, if (3.17) holds for v, then its projection into the space $\mathcal{H}_N = \text{span}\{K_{x_i}, \cdots, K_{x_N}\}$, written as v_N, solves the set of equations as well, since $\langle v - v_N, K_{x_i}\rangle = 0$ for any $i \in \{1, \cdots, N\}$. Furthermore, $\|v_N\| \leqslant \|v\|$, thus the minimum norm solution, if exists at all, lives in $\mathcal{H}_N$. Under assumption (3.8), the matrix $A = (K(x_i, x_j))$ is invertible and the unique minimum solution is given by

$$(a_1, \cdots, a_N)A^{-T}(K_{x_1}, \cdots, K_{x_N})^T \tag{3.18}$$

Our next theorem gives the kernel condition under which kernel integral expression can be used to approximate an arbitrary L^1 integrable function. I also refer to the coming work [20] for a generalization of this result to higher dimension situation.

Theorem 3.3. *Given $K \in L^1([-1, 1])$ and a dense subset $\Gamma = \{\sigma_i\}_i$ of the rotation group, the span of functions $\{\mathcal{R}(\sigma_i)K(\xi_0 \cdot y)\}$ is dense in $L^1(\mathbb{S}^2)$ if and only if $\hat{K}_l \neq 0$ for all $l \in \mathbb{N}$.*

Proof. Firstly observe that for any $g \in L^\infty(\mathbb{S}^2)$ there is

$$\begin{aligned} \widehat{(K * g)}_l^m &= \int_{\mathbb{S}^2}\int_{\mathbb{S}^2}\int_{\mathbb{S}^2} K(x \cdot y)P_l(x \cdot z)g(y)dxdyY_l^m(z)dz \\ &= \hat{K}_l \int_{\mathbb{S}^2}\int_{\mathbb{S}^2} P_l(y \cdot z)Y_l^m(z)g(y)dzdy \\ &= \hat{K}_l\hat{g}_l^m \end{aligned} \tag{3.19}$$

hence it is easy to see that $K_l \neq 0$ for all $l \in \mathbb{N}$, if and only if the equation

$$K * g(x) = 0 \quad a.e. \tag{3.20}$$

does not have any bounded solution other than the zero function.

Let $V = \{$all finite linear combinations $\sum_i c_i K(x_i \cdot y)\}$ and denote by $\bar{V}$ its closure. For every $S \in \bar{V}$ and $\epsilon > 0$, suppose $\|S - S_N\|_{L^1} \leqslant \frac{1}{N}$ with $S_N(y) = \sum_i c_i K(x_i \cdot y)$, we see immediately from the fact that

$$\int_{\mathbb{S}^2} |(S - S_N)(y) g(y)| dy \leqslant \|g\|_\infty \cdot \frac{1}{N}$$

the equation (3.20) holds if and only if

$$\int_{\mathbb{S}^2} S(y) g(y) dy = 0 \tag{3.21}$$

for any $S \in \bar{V}$, which is further equivalent to the condition that $\bar{V} = L^1(\mathbb{S}^2)$. Indeed, Hahn-Banach Theorem implies that if $\bar{V} \neq L^1(\mathbb{S}^2)$, we can find a nonzero continuous linear functional v in $L^1(\mathbb{S}^2)$ which is vanishing on $\bar{V}$. Since $\mathbb{S}^2$ is of finite measure, by Riesz-representation Theorem there exists a bounded nonzero function g such that $v(S) = \int_{\mathbb{S}^2} S(y) g(y) dy = 0$ for any $S \in \bar{V}$, a contradiction. Thus we have proved that the span of all rotated functions $\{\mathcal{R}(\sigma_i) K(\xi_0 \cdot y)\}$ being dense in $L^1(\mathbb{S}^2)$ is equivalent to the condition $\hat{K}_l \neq 0$ for $\forall l \in \mathbb{N}$.

Finally, by the density of continuous functions in L^1 space, for any $f \in L^1$ there exist $K^c \in C([-1,1])$ and some set $\{\sigma_i^N\}_i$ such that $\|S_N - S_N^c\|_{L^1} < \frac{1}{N}$ and $\|S_N - f\| < \frac{1}{N}$, where $S_N^c = \sum_i c_i \mathcal{R}(\sigma_i^N) K^c(\eta_0 \cdot y)$. Now by the density assumption of Γ in $SO(3)$, for arbitrary $\epsilon > 0$, we can find some $\{\sigma_{i'}\} \subset \Gamma$ such that $\|f - \sum_{i'} c_{i'} \mathcal{R}(\sigma_{i'}) K^c(\eta_0 \cdot y)\|_{L^1} < \epsilon$, hence the claimed result follows. $\square$

Similarly, for a family of kernels $\{K^j\}_{j \in J}$ we have the following extension

Corollary 3.4. *The equations*

$$\int_{\mathbb{S}^2} K^j(x \cdot y) g(y) dy = 0 \ \forall x \in \mathbb{S}^2, j \in J$$

have a nonzero bounded solution g when and only when for some $l \in \mathbb{N}$, $\hat{K}_l^j = 0$ for all $j \in J$.

Proof. If $\hat{K}_l^j = 0$ for all $j \in J$, then $\int_{\mathbb{S}^2} K^j(x \cdot y) Y_l^m(y) dy = 0$ for all x and j. Conversely, if $\int_{\mathbb{S}^2} K^j(x \cdot y) g(y) dy = 0$ has a nonzero bounded solution g for all j, then $0 = Z_l * (K^j * g) = \hat{K}_l^j \mathbb{P}_l g$. Since g is not identically zero, there exists some $\mathbb{P}_l g \neq 0$, hence $\hat{K}_l^j = 0$ for all $j \in J$. $\square$

Remark 3.5. In the proof we already use the fact that $g \in L^\infty(\mathbb{S}^2) \subset L^2(\mathbb{S}^2)$, otherwise (3.19) does not make sense. In fact, the theorem is still valid even if we set $g \in L^2(\mathbb{S}^2)$, and the corresponding proof is not much changed.

Let ν be an arbitrary linear continuous functional on $L^1([-1,1])$. Due to the fact that $\nu(K) = \int_{-1}^{1} K(t)g_\nu(t)dt$ for some $g_\nu \in L^\infty([-1,1])$ and that $\|\nu\| = \sup\limits_{\|g_\nu\|_\infty=1} \int_{-1}^{1} K(t)g_\nu(t)dt$, as an immediate consequence of theorem 3.3 is following expansion result for linear continuous functionals on $L^1(\mathbb{S}^2)$ in terms of translations of ν.

Corollary 3.6. *Given a dense subset* $\Gamma = \{\sigma_i\}_i$ *of the rotation group and* $\nu \in (L^1(\mathbb{S}^2))^*$ *such that* $\ker \nu = \{0\}$*, there is*

$$(L^1(\mathbb{S}^2))^* = \overline{span\{\nu_i\}}$$

where

$$\nu_i(f) = \int_{\mathbb{S}^2} f(y)\mathcal{R}(\sigma_i)g_\nu(\eta_0 \cdot y)dy$$

for any $f \in L^1(\mathbb{S}^2)$.

Proof. We only need to prove $(L^1(\mathbb{S}^2))^* \subset \overline{\text{span}\{\nu_i\}}$. For any $\mu \in (L^1(\mathbb{S}^2))^*$ and $f \in L^1(\mathbb{S}^2)$, there exists $h \in L^\infty(\mathbb{S}^2)$, such that $\mu(f) = \int_{\mathbb{S}^2} f(y)h(y)dy$. Due to our assumption that $\ker \nu = \{0\}$, (3.20) implies that $\widehat{g_\nu}_l = 0$ for any l. Consequently, given any $\epsilon > 0$ there exist some c_i and $\{\sigma_i\}_i \subset \Gamma$ such that $\|h(y) - \sum\limits_i c_i\mathcal{R}(\sigma_i)g_\nu(\eta_0 \cdot y)\|_{L^1(\mathbb{S}^2)} < \epsilon$, hence we have

$$\left|\mu(f) - \sum_i c_i\nu_i(f)\right| < \epsilon\|f\|_\infty$$

for any $f \in L^\infty(\mathbb{S}^2)$ such that $\|f\|_{L^1} = 1$, since L^∞ is dense in L^1 on compact set. The claim follows from the arbitrariness of ϵ. ☐

Let $\mathcal{Q}_a * g := \int_{\mathbb{S}} K(x \cdot y_a)g(y)dy$, the observation that

$$\begin{aligned} \widehat{\mathcal{Q}_a * g}(l,m) &= \int\int K(x \cdot y_a)Y_l^m(x)g(y)dydx \\ &= \int \hat{K}(l)Y_l^m(y_a)g(y)dy \end{aligned} \tag{3.22}$$

leads to

$$\lim_{a\to 0^+} \widehat{\mathcal{Q}_a * g}(l,m) = \frac{1}{\sqrt{4\pi}}\hat{K}(l)\delta_{m,0}\int_{\mathbb{S}^2} g(x)dx$$

or

$$\lim_{a\to 0^+} \mathcal{Q}_a * g(y) = \sum_l \frac{\sqrt{2l+1}}{4\pi} \hat{K}(l) \int_{\mathbb{S}^2} g(x) dx P_l(\cos\theta_y) \tag{3.23}$$

and

$$\lim_{a\to 1} \widehat{\mathcal{Q}_a * g}(l,m) = \hat{K}(l)\hat{g}_{lm}$$

or equivalently

$$\lim_{a\to 1} \int_{\mathbb{S}^2} |\mathcal{Q}_a * g(y) - g(y)|^2 dy = \sum_l \sum_{|m|\leqslant l} |\hat{K}(l) - 1|^2 |\hat{g}_{lm}|^2 \tag{3.24}$$

Proposition 3.7. *$\{\mathcal{Q}_a\}$ is an approximate identity at $a = 1$ in L^2 iff $\hat{K}(l) = 1$ for any $l \in \mathbb{N}$.*

Let I be the parameter set of frames $\{\phi_i\}_{i\in I}$ and $\{\phi_i^\dagger\}_{i\in I}$. Consider kernels of the form

$$K^t(x,y) = \sum_{i\in I} \lambda_i(t)\phi_i(x)\phi_i^\dagger(y) \tag{3.25}$$

with $\lambda_i \in \ell^\infty$ such that $K^t(x,y)$ exist for all t, x, y.

Definition 3.8. Let us call the kernel in (3.25) a *frame kernel.* Set $I_N \subset I$ with $\|I_N\|_0 = N$ and $I_N^\perp = I \setminus I_N$. Denote by $K_{I_N}^t(x,y) = \sum_{i\in I_N} \lambda_i(t)\phi_i(x)\phi_i^\dagger(y)$. If $\langle f, K^t(x,\cdot)\rangle = f(x)$, then $f_{I_N}(x) = \langle f(\cdot), K_{I_N}(x,\cdot)\rangle$ is called a *N-term kernel approximation* of f. When

$$\|\langle f(\cdot), K_{I_N^\perp}^t(\cdot,y)\rangle\| \leqslant \|\langle f(\cdot), K_{J_N^\perp}^t(\cdot,y)\rangle\|$$

for any $J_N \subset I$ with $\|J_N\|_o = N$ we call f_{I_N} a *best N-term kernel approximation* of f.

Remark 3.9. In contrast to the standard text in approximation theory where N-term approximation is usually reserved for nonlinear spaces, here I do not distinguish between linear and nonlinear spaces.

Example 3.10. Clearly frame property (1.2) can be reformulated in the kernel means

$$A\|f\|^2 \leqslant \langle f(x), \langle K^t(x,y), f(y)\rangle\rangle \leqslant B\|f\|^2 \tag{3.26}$$

in the special case that $\lambda_i = 1$ for all $i \in I$ and $\phi_i = \phi_i^\dagger$.

If $\{\phi_i^\dagger\}_{i\in I}$ is the dual frame of $\{\phi_i\}_{i\in I}$ and $\lambda_i = 1$, then

$$\|f\|^2 = \langle f(x), \langle K^t(x,y), f(y)\rangle\rangle \tag{3.27}$$

namely K^t is a reproducing kernel.

If $\lambda_i(t) = e^{-c_i t}$ with c_i the eigenvalues of Laplace-Beltrami operator on a manifold, we call

$$h_t(x,y) = \sum_{i=0}^{\infty} e^{-c_i t}\phi_i(x)\phi_i^\dagger(y) \tag{3.28}$$

a *generalized heat kernel.*

Another good example of frame kernel is the multiscale kernel we have defined in (3.6), its reproducing property is given in Theorem 3.13. In addition, if we denote by $A^\dagger$ and $B^\dagger$ the lower and upper bound of frame $\{\phi_i^\dagger\}_{i\in I}$ respectively, from the definition it is obvious that

$$\begin{aligned} \|\langle f(\cdot), K_{I_N^\perp}(\cdot,y)\rangle\| &= \sup_{\|g\|=1}\left|\langle\langle f(\cdot), K_{I_N^\perp}(\cdot,y)\rangle, g(y)\rangle\right| \\ &= \sup_{\|g\|=1}\|\{\lambda_i\langle f,\phi_i\rangle\}_{i\in I_N^\perp}\|_{\ell^2}\|\{\langle\phi_i^\dagger, g\rangle\}_{i\in I_N^\perp}\|_{\ell^2} \\ &\leqslant \sqrt{BB^\dagger}\max_i|\lambda_i|\|f\| \end{aligned} \tag{3.29}$$

and by assuming without loss of generality that $\lambda_i \neq 0$, there is

$$\|\langle f(\cdot), K_{I_N^\perp}(\cdot,y)\rangle\| \geqslant \sqrt{AA^\dagger}\min_i|\lambda_i|\|f\| \tag{3.30}$$

The error of the best l-term Legendre kernel approximation of image with smooth boundary decays as $l^{-1/2}$ when $l \to \infty$, which is shown in the next proposition by choosing the characteristic function of the spherical cap $C(\xi_0,\Theta)$ centered at ξ_0. The following asymptotic result about Legendre polynomials can be derived from our Lemma 2.14 and can also be found in [18] and [71].

Lemma 3.11. *Let c be a fixed positive constant, $l \to \infty$. Then*

$$P_l^{(m)}(\cos\theta) = \begin{cases} \theta^{-m-\frac{1}{2}}O(l^{m-\frac{1}{2}}) & \text{if } c/l \leq \theta \leq \pi/2 \\ O(l^{2m}) & \text{if } 0 \leq \theta \leq c/l \end{cases} \tag{3.31}$$

In fact, the first term of the asymptotic expansion of $P_l^m(\cos\theta)$ is

$$P_l^m(\cos\theta) = (-l)^m\left(\frac{2}{\pi l\sin\theta}\right)^{\frac{1}{2}}\cos[(l+\frac{1}{2})\theta + \frac{(m-1)}{2}\pi] + O(l^{\frac{-3}{2}}) \tag{3.32}$$

where $P_l^m(\cos\theta) = (-1)^l\sin^m\theta P_l^{(m)}(\cos\theta)$, $\epsilon \leq \theta \leq \pi-\epsilon$, $\epsilon > 0$, $l \gg m$, $l \gg \frac{1}{\epsilon}$.

Proposition 3.12. $\min_{I_L} \| \sum_{l \notin I_L} \mathbb{P}_l \chi_{C(\xi,\Theta)} \|_2 = O(L^{-1/2})$, *namely the error of best N-term kernel approximation with respect to $\{P_l\}_{l \geqslant 0}$ decays at least as $O(N^{-1/2})$;*

Proof. Let $f = \chi_{C(\xi,\Theta)}$, then

$$\langle f, P_l(\omega \cdot) \rangle = b(l, \Theta) P_l(\omega \cdot \xi) \tag{3.33}$$

where $b(l, \Theta) = 2\pi \int_0^\Theta P_l(\cos\theta) \sin\theta d\theta$.

By integrating (4.44) we obtain

$$\begin{aligned} \int_0^\Theta P_l(\cos\theta) \sin\theta d\theta &= (l+1)^{-1} \left(P_{l-1}(\cos\Theta) \right) - \cos\Theta P_l(\cos\Theta)) \\ &= O(l^{-3/2}) \end{aligned} \tag{3.34}$$

By (3.31), if $\xi \cdot \omega \neq 1$, $P_l(\omega \cdot \xi) = O(l^{-1/2})$ as $l \to \infty$; If $\xi \cdot \omega = 1$, $P_l(\omega \cdot \xi) = O(1)$. Therefore

$$\begin{aligned} \|\langle \chi_{(\xi,\Theta)}, Z_l(\omega \cdot) \rangle\|_2^2 &= \Big(\int_{|\omega \cdot \xi - 1| \geqslant \frac{1}{l}} + \int_{|\omega \cdot \xi - 1| \leqslant \frac{1}{l}} \Big) |b(l,\Theta) Z_l(\omega \cdot \xi)|^2 d\Omega(\omega) \\ &= O(l^{-2}) + O(l^{-1} \cdot l^{-1}) \\ &= O(l^{-2}) \end{aligned}$$

due to the observation that surface area of the region $\{\omega : |\omega \cdot \xi - 1| \leqslant \frac{1}{l}\}$ is of size $O(l^{-1})$.

$$\sum_L^\infty \|\mathbb{P}_l f\|_2^2 = \sum_L^\infty \|\langle \chi_{(\xi,\Theta)}, Z_l(\omega \cdot) \rangle\|_2^2 = O(L^{-1}) \tag{3.35}$$

Thus the error of best L-term kernel approximation is

$$\|f - f_L\|_2 = O(L^{-1/2}) \tag{3.36}$$

□

As I promised, now let us turn our attention back to the kernel $K_{\Phi,\lambda}$ defined in (3.6). Endow the space V_j with norms $\|f * \Phi_j * \Phi_j\|_{V_j}^2 = \sum_l \left| \hat{f}(l) \right|^2$, we can claim the following result.

Theorem 3.13. *Given γ satisfying the conditions in* (2.5) *and a positive sequence $\lambda = \{\lambda_{j,l}\}_{j,l}$ such that $\sum\limits_{j,l} \lambda_{j,l} < \infty$ and $\mu_j = \gamma(2^{-j}l)^4/\lambda_{j,l}$ independent of l, then $K_{\Phi,\lambda}(x,y)$ forms a reproducing kernel of the Hilbert space*

$$\mathscr{H}_{\Phi,\lambda} = \{g \in L^2 : g = \sum_j g_j \in \bigcup_j V_j \text{ and } \sum_j \mu_j \|g_j\|_{V_j}^2 < \infty\} \tag{3.37}$$

with norm

$$\|g\|_{\Phi,\lambda}^2 = \inf\{\sum_j \mu_j \|g_j\|_{V_j}^2 : g = \sum_j g_j\} \tag{3.38}$$

Proof. Recall that $\widehat{\Phi}_j(l) = \gamma(2^{-j}l)$ and let us define

$$\mathsf{A}^*_{\{\Phi_j\}_{j\geqslant 0}}(\mathbf{c}) = \sum_{l,j\geqslant 0} c_{l,j}\widehat{\Phi}_j(l)^2 P_l(x\cdot\xi_0) \tag{3.39}$$

or abbreviated as A^* without confusion, where

$$\mathbf{c} \in \ell^2_\mu = \{\{c_{l,j}\}_{l,j} : \sum_{l,j} \mu_j |c_{l,j}|^2 < \infty\} \tag{3.40}$$

Denote by $N_{\mathsf{A}^*} = \{\mathbf{c} \in \ell^2_\mu : \mathsf{A}^*(\mathbf{c}) = 0\}$ the null space of A^*. With the inner product

$$\left\langle \sum_{l,j} c_{l,j}\widehat{\Phi}_j(l)^2 P_l(x\cdot\xi_0), \sum_{l,j} c'_{l,j}\widehat{\Phi}_j(l)^2 P_l(x\cdot\xi_0)\right\rangle_{\mathsf{A}^*} = \left\langle P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}), P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}')\right\rangle_{\ell^2_\mu} \tag{3.41}$$

the range of A^* becomes a Hilbert space, where $P_{N^\perp_{\mathsf{A}^*}}$ is the projection onto the sequence subspace $N^\perp_{\mathsf{A}^*}$. Indeed, it is obviously bilinear and symmetric. It is well-defined, for $\mathsf{A}^*(\mathbf{c}) = \mathsf{A}^*(\mathbf{d})$ implying that $P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}) = P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}-\mathbf{d}+\mathbf{d}) = P_{N^\perp_{\mathsf{A}^*}}(\mathbf{d})$. Furthermore, it is clear that $\langle \mathsf{A}^*(\mathbf{c}), \mathsf{A}^*(\mathbf{c})\rangle \geqslant 0$, and equality holds when and only when $P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}) = \mathbf{0}$. The completeness of RanA^* follows from that of ℓ^2_λ, due to the fact that A^* is an isometric isomorphism.

Fix x on the sphere. $\{\frac{1}{\mu_j}\gamma(2^{-j}l)^2 P_l(x\cdot\xi_0)\}_{j,l}$ forms a sequence in ℓ^2_μ under our assumption, hence $K_{\Phi,\lambda}(x,\cdot) \in$ RanA^*. Notice that

$$\begin{aligned}
&\langle \mathsf{A}^*(\mathbf{c}), K_{\Phi,\lambda}(x,\cdot)\rangle_{\mathsf{A}^*} \\
=&\left\langle \sum_{l,j\geqslant 0} c_{l,j}\widehat{\Phi}_j(l)^2 P_l(y\cdot\xi_0), \sum_{l,j\geqslant 0}\frac{1}{\mu_j}\widehat{\Phi}^4_j(l) P_l(x\cdot\xi_0)P_l(y\cdot\xi_0)\right\rangle_{\mathsf{A}^*} \\
=&\sum_{l,j} c_{l,j}\widehat{\Phi}^2_j(l)P_l(x\cdot\xi_0)
\end{aligned}$$

Thus $K_{\Phi,\lambda}$ is a reproducing kernel for the Hilbert space $(\mathrm{Ran}\mathsf{A}^*, \langle\cdot,\cdot\rangle_{\mathsf{A}^*})$. Furthermore, since for any $\mathbf{c}' \in \ell^2_\mu$ such that $\mathsf{A}^*(\mathbf{c}') = \mathsf{A}^*(P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}))$ there is $\|P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c})\|^2_{\ell^2_\mu} = \|P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c}')\|^2_{\ell^2_\mu} \leqslant \|\mathbf{c}'\|^2_{\ell^2_\mu}$, it follows that

$$\begin{aligned}\|\mathsf{A}^*(\mathbf{c})\|^2_{\mathsf{A}^*} &= \|P_{N^\perp_{\mathsf{A}^*}}(\mathbf{c})\|^2_{\ell^2_\mu} \\ &= \min\{\sum_{l,j}\mu_j|c'_{l,j}|^2 : \mathbf{c}' \in \ell^2_\mu, \mathsf{A}^*(\mathbf{c}') = \mathsf{A}^*(\mathbf{c})\} \\ &= \|\mathsf{A}^*(\mathbf{c})\|^2_\Phi\end{aligned}$$

hence the claim of the theorem holds. □

3.2 Construction of spherical frames

The goal of this section is to establish frame properties for two types of representation systems on the sphere that we have met in Chapter 2. Recall that the first type considers dilation in the frequency domain. For this type, starting with a general basis, for instance the spherical harmonics, our next theorem allows us to transform them into a new frame suited to the multiresolution structure and at the same time equips them with good properties such as local support and fast decay. Interestingly, as far as I know when I am writing this part, it has never been discussed by any before. For that purpose, we firstly need a lemma, whose proof can be found in [81] for instance, hence omitted here.

Lemma 3.14. *If $\{\phi_k\}_{k\geqslant 0}$ is a frame for Hilbert space $\mathscr{H}$ with bounds A and B, and T is an bounded operator on $\mathscr{H}$ with closed range, then $\{T\phi_k\}_{k\geqslant 0}$ is a frame for the range of T with bounds $A\|\bar{T}\|^{-2}$ and $B\|T\|^2$, where $\bar{T}$ is the pseudo-inverse of T.*

Let $\Xi \subset \mathbb{N}$ and $\mathcal{H}_\Xi = \bigoplus_{l\in\Xi}\mathcal{H}_l$ be a subspace of $L^2(\mathbb{S}^2)$ and denote by $V_j = \Phi_j * \Phi_j * \mathcal{H}_\Xi$ and $W_j = \Psi_j * \tilde{\Psi}_j * \mathcal{H}_\Xi$ the spaces of multiresolution for the subspace $\mathcal{H}_\Xi$.

Theorem 3.15. *Let γ be piecewise differentiable and admissible in the sense of (2.5) with $\inf_{l\in\Xi}|\gamma(l)| > 0$. Given a frame $\{b_k\}$ for $\mathcal{H}_\Xi$, then $\{\phi_{k,j} = \frac{1}{(\sqrt{2}c)^j}\Phi_j * \Phi_j * b_k\}$ is a frame for V_j. If additionally for some constant $\tau > 0$, $t_0 \geqslant 0$ there is*

$$-\gamma'(t) \geqslant \frac{\tau\gamma(t)}{t} \quad \text{for all } t > t_0 \tag{3.42}$$

then $\{\psi_{k,j} = \frac{1}{(\sqrt{2}c)^j}\Psi_j * \tilde{\Psi}_j * b_k\}$ is a frame for $W_j \cap \mathcal{H}_{\Xi[t_0,j]}$, where $\Xi[t_0,j]$ is any subset of $\{l : l \geqslant t_0 2^{j+1}\}$.

Proof. Suppose $\{b_k\}$ is a frame with lower bounds A and upper bounds B. Define a new function class

$$\mathcal{G}_{2,c} = \{f \in L^2(\mathbb{S}^{d-1}) : \|\mathbb{P}_l f\| \leqslant c\|\mathbb{P}_{l'} f\| \text{ for any } l \geqslant 2l^{'}\} \tag{3.43}$$

where $c > 0$. Notice that $\{\phi_{k,0}\}$ forms a frame for V_0. For any $f \in \mathcal{H}_\Xi$ and $g = \Phi * \Phi * f \in V_0$ there is

$$\sum_k |\langle g, \phi_{k,0}\rangle|^2 \leqslant B\| \sum_l \gamma^4(l)\mathbb{P}_l f\|^2 \leqslant \sup_l \gamma^2(l)\|g\|^2 B$$

and similarly $\sum\limits_k |\langle g, \phi_{k,0}\rangle|^2 \geqslant \inf\limits_{l\in\Xi} \gamma^2(l)\|g\|^2 A > 0$.

Introduce the operator

$$D_j g = \frac{1}{(\sqrt{2}c)^j} \sum_{l\in\Xi} \gamma(2^{-j}l)^2 \mathbb{P}_l f \tag{3.44}$$

on $V_0 \cap \mathcal{G}_{2,c}$.

It is clearly a map onto $V_{j,c} = \Phi_j * \Phi_j * \mathcal{H}_\Xi \cap \mathcal{G}_{2,c}$, which consists solely of functions like $\Phi_j * \Phi_j * f$. Besides, by separating the sum into even terms and odd terms, under our assumption that $f \in \mathcal{G}_{2,c}$ we can get an estimation

$$\begin{aligned}
\|D_{j+1}g\|^2 &= (2c^2)^{-j-1}\left[\sum_{\text{odd}} \gamma^4(2^{-j-1}l)\|\mathbb{P}_l f\|^2 + \sum_{\text{even}} \gamma^4(2^{-j-1}l)\|\mathbb{P}_l f\|^2\right] \\
&\leqslant (2c^2)^{-j-1} \sum_{l\geqslant 0} \gamma^4(2^{-j}(l+\frac{1}{2}))\|\mathbb{P}_{2l+1} f\|^2 \\
&+ (2c^2)^{-j-1} \sum_{l\geqslant 0} \gamma^4(2^{-j}l)\|\mathbb{P}_{2l} f\|^2 \\
&\leqslant \|D_j g\|^2
\end{aligned}$$

for $j \geqslant 0$, hence

$$(\sqrt{2}c)^{-j} I \leqslant D_j \leqslant I \tag{3.45}$$

and

$$I \leqslant \overline{D}_j = D_{-j} \leqslant (\sqrt{2}c)^j I \tag{3.46}$$

Meanwhile $D_j - \sqrt{2}cD_{j+1}$ is an operator onto $W_{j,c}$, and there is the estimation

$$\begin{aligned}
\left\|\left(D_j - \sqrt{2}cD_{j+1}\right) g\right\|_2^2 &= (2c^2)^{-j} \sum_l \left(\gamma(2^{-j}l)^2 - \gamma(2^{-j-1}l)^2\right)^2 \|\mathbb{P}_l f\|^2 \\
&\leqslant 2c^2 \|D_{j+1} g\|_2^2
\end{aligned} \tag{3.47}$$

and for $g \in W_{j,c} \cap \mathcal{H}_{\Xi[t_0,j]}$

$$\begin{aligned}
&\left\|\overline{\left(D_j - \sqrt{2}cD_{j+1}\right)g}\right\|_2^{-2} \\
&= \left\|\left(D_j - \sqrt{2}cD_{j+1}\right) g\right\|_2^2 \\
&= (2c^2)^{-j} \sum_{l \geqslant t_0 2^{j+1}} \left|\int_{2^{-j-1}l}^{2^{-j}l} -\gamma^{'}(t)dt\right|^2 \left(\gamma(2^{-j}l) + \gamma(2^{-j-1}l)\right)^2 \|\mathbb{P}_l f\|^2 \qquad (3.48) \\
&\geqslant (2c^2)^{-j} \sum_{l \geqslant t_0 2^{j+1}} \gamma^2(2^{-j}l) \left|\int_{2^{-j-1}l}^{2^{-j}l} \frac{\tau}{t} dt\right|^2 \cdot 4\gamma^2(2^{-j}l)\|\mathbb{P}_l f\|^2 \\
&= 4\tau^2 \ln^2 2\|D_j g\|^2
\end{aligned}$$

due to (3.42).

Those together with Lemma 3.14 imply that $\{\phi_{l,j}\}_{l\geqslant 0}$ form a frame for $V_{j,c}$, while $\{(D_j - D_{j+1})\, \phi_{l,0}\}_{l\geqslant 0}$ form a frame for the space $W_{j,c} \cap \mathcal{H}_{\Xi[t_0,j]} = \Phi_j * \Phi_j * \mathcal{H}_\Xi \cap \mathcal{G}_{2,c} \cap \mathcal{H}_{\Xi[t_0,j]}$.

Taking any $h \in L^2(\mathbb{S}^{d-1})$ and $c > 1$, suppose $\bar{l}_1$ is the smallest degree such that $\mathbb{P}_{\bar{l}_1} h \neq 0$. Let $h_{\bar{l}_0}(x) = 0$ and

$$u_{\bar{l}_0}(x) = \sum_{l=0}^{\min\{\bar{l}_1-1,0\}} c^{-l} \sqrt{\frac{4\pi}{(2l+1)}} Z_l(x \cdot \xi_0) \qquad (3.49)$$

Define inductively, for the smallest degree $\bar{l}_{k+1} \geqslant \bar{l}_k$ such that $\mathbb{P}_{\bar{l}_{k+1}} h \neq 0$, the functions

$$h_{\bar{l}_{k+1}} = h_{\bar{l}_k} + u_{\bar{l}_k} + \frac{\|h_{\bar{l}_k} - h_{\bar{l}_{k-1}} - u_{\bar{l}_{k-1}}\|}{c^{\bar{l}_{k+1}-\bar{l}_k}} \frac{\mathbb{P}_{\bar{l}_{k+1}} h}{\|\mathbb{P}_{\bar{l}_{k+1} h}\|} \qquad (3.50)$$

and

$$u_{\bar{l}_{k+1}} = \sum_{l=\bar{l}_k+1}^{\bar{l}_{k+1}-1} c^{-l+\bar{l}_k} \sqrt{\frac{4\pi}{(2l+1)}} Z_l(x \cdot \xi_0) \|h_{\bar{l}_{k+1}} - h_{\bar{l}_k} - u_{\bar{l}_k}\| \qquad (3.51)$$

Then $h \in \overline{\operatorname{span}\{h_{\bar{l}_{k+1}} - (h_{\bar{l}_k} + u_{\bar{l}_k}) : k \geqslant 0\}} \subset \overline{\mathcal{G}_{2,c}}$, namely $\mathcal{G}_{2,c}$ is dense in $L^2(\mathbb{S}^{d-1})$. Consequently $V_{j,c}$ is dense in V_j and $W_{j,c}$ is dense in W_j, hence $\{\phi_{k,j}\}$ and $\{\psi_{k,j}\}$ are frames of V_j and $W_j \cap \mathcal{H}_{\Xi[t_0,j]}$ with the same bounds respectively.

□

Corollary 3.16. *Under the same assumptions of theorem* 3.15 *with* (3.42) *being replaced by*

$$\gamma(2^{-1}t) - \gamma(t) \geqslant \tau\gamma(t) \text{ for all } t > t_0 \tag{3.52}$$

$\{\psi_{k,j}\}$ *form a frame for* $W_j \cap \mathcal{H}_{\Xi[t_0,j-1]}$.

Proof. This claim follows from the same line proof of the above theorem except that here we have for $g \in W_j \cap \mathcal{H}_{\Xi[t_0,j-1]}$ that

$$\begin{aligned}
&\left\| \left(D_j - \sqrt{2}cD_{j+1} \right) g \right\|_2^2 \\
&= (2c^2)^{-j} \sum_{l \geqslant t_0 2^j} \left| \gamma(2^{-j-1}l) - \gamma(2^{-j}l) \right|^2 \left(\gamma(2^{-j}l) + \gamma(2^{-j-1}l) \right)^2 \|\mathbb{P}_l f\|^2 \\
&\geqslant (2c^2)^{-j}\tau^2 \sum_{l \geqslant t_0 2^j} \gamma^2(2^{-j}l) \cdot 4\gamma^2(2^{-j}l)\|\mathbb{P}_l f\|^2 \\
&= 4\tau^2 \|D_j g\|^2
\end{aligned} \tag{3.53}$$

□

Example 3.17. The Shannon type wavelets is a good example for our theory here. Let

$$\widehat{\Phi}_j(l) = \begin{cases} 1 & l \leqslant 2^j - 1 \\ 0 & l \geqslant 2^j \end{cases} \tag{3.54}$$

with corresponding wavelets

$$\widehat{\Psi}_j(l) = \widehat{\widetilde{\Psi}}_j(l) = \begin{cases} 1 & 2^j \leqslant l \leqslant 2^{j+1} - 1 \\ 0 & \text{else} \end{cases} \tag{3.55}$$

It is clear that $W_j \perp V_j$ and $V_j \bigoplus W_j = V_{j+1}$ in this situation. Thus $W_i \perp W_j$ for any $i \neq j$. By choosing $t_0 = 1$, $\mathcal{H}_{\Xi[1,j-1]} = \{l : l \geqslant 2^j\}$, we see that for

$$\gamma(t) = \begin{cases} 1 & t \in [0,1) \\ 0 & t \geqslant 1 \end{cases}$$

inequality $\gamma(\frac{t}{2}) - \gamma(t) \geqslant \gamma(t)$ holds for any $t > 1$, hence by Corollary 3.16 $\bigcup_{j=0}^{J} \{\psi_{k,j} : k \geqslant 0\}$ form a wavelet frame for $\bigoplus_{j=0}^{J} W_j \cap \mathcal{H}_{\Xi[1,j-1]} = \bigoplus_{j=0}^{J} W_j$.

The second type spherical wavelets we have encountered is constructed through stereographic projection. In proposition 2.18 it has been proved that a plane wavelet under certain conditions gives an admissible wavelet on the sphere.

Since stereographic projection π is conformal, hence preserves angle, a plane wavelet frame $\{\psi_i\}_{i\in I}$ of unit norm with lower and upper frame bounds A,B respectively immediately implies that for any $f \in L^2(\mathbb{S}^2)$ there is

$$AC\|f\|^2_{L^2(\mathbb{S}^2)} \leqslant \sum_{i\in I}\langle \pi^{-1}\pi f, \pi^{-1}\psi_i\rangle_{L^2(\mathbb{S}^2)} \leqslant BC\|f\|^2_{L^2(\mathbb{S}^2)} \tag{3.56}$$

where $C = \|\pi f\|^2_{L^2(\mathbb{R}^2)}/\|f\|^2_{L^2(\mathbb{S}^2)}$, namely $\{\pi^{-1}\psi_i\}_{i\in I}$ is automatically a frame on the sphere. It seems that nothing needs to be done further, due to all kinds of plane wavelets with good frame properties having been well studied. However, a closer look tells a different story. Since the natural lattice on the plane does not generate a good grid on the sphere through projection, it is necessary to establish principles or conditions, under which we can obtain spherical frames directly on the sphere. This is the next theorems in this section about.

I would like to mention that main differences between our second type frame construction and those kernel based frames in [63] that apply to more general Lie groups, include but not limited to: the latter is solely designed for isotropic case while ours includes anisotropic case; and the latter starts from continuous wavelets without a result like Theorem 3.19 but rather directly utilizes the Calderón reproducing formula with respect to the scaling parameter, hence immediately having reproducing property, while in the system we introduced above, Calderón's formula does not apply. However, in this work we are able to achieve tight frames for continuous and discrete α-wavelets/shearlets on the sphere in a different way.

To meet our purpose let us firstly extend the traditional spherical harmonics, as a natural generalization of the solutions of (4.41).

Definition 3.18. Let the *spherical harmonics of fractional degree* λ *and order* β to be

$$Y^\beta_\lambda(\theta,\varphi) = c_{\lambda,\beta}e^{i\beta\varphi}P^\beta_\lambda(\cos\theta) \tag{3.57}$$

where $\lambda, \beta \in \mathbb{R}$, $c_{\lambda,\beta} = e^{i\beta\pi}\sqrt{\frac{(2\lambda+1)\Gamma(\lambda-\beta+1)}{4\pi\Gamma(\lambda+\beta+1)}}$ and P^β_λ are associated Legendre functions which solve the differential equation

$$(1-z^2)u^{''} - 2zu^{'} + \left[\lambda(\lambda+1) - \frac{\beta}{1-z^2}\right]u = 0 \tag{3.58}$$

and such that $\int_{-1}^1 |P^\beta_\lambda(t)|^2 dt = \frac{2\Gamma(\lambda+\beta+1)}{(2\lambda+1)\Gamma(\lambda-\beta+1)}$.

For $|z-1| < 2$ in the complex domain, there is the following expression(see for instance [11])

$$P^\beta_\lambda(z) = d_{\lambda,\beta}\ {}_2F_1(\lambda+\beta+1, -\lambda+\beta, \beta+1; \frac{1-z}{2})(z^2-1)^{\frac{\beta}{2}} \tag{3.59}$$

where ${}_2F_1(a,b,c;z)$ is a solution of the hypergeometric equation

$$z(1-z)\frac{d^2u}{dz^2}+[c-(a+b+1)z]\frac{du}{dz}-abu=0 \tag{3.60}$$

and

$$d_{\lambda,\beta}=\frac{(-1)^\beta}{2^\beta\Gamma(\beta+1)}\frac{\Gamma(\lambda+\beta+1)}{\Gamma(\lambda+1)}(-\lambda+\beta-1)(-\lambda+\beta-2)\cdots(-\lambda) \tag{3.61}$$

when $\lambda>0$, $-\lambda+\beta$ and $-\lambda-\beta$ are negative integers; by identity $\Gamma(x)\Gamma(1-x)=\frac{\pi}{\sin(\pi x)}$ for non-integer values x, there is

$$\begin{aligned} d_{\lambda,\beta}&=\frac{(-1)^\beta}{2^\beta\Gamma(\beta+1)}\frac{\Gamma(\lambda+\beta+1)\Gamma(-\lambda+\beta)}{\Gamma(\lambda+1)\Gamma(-\lambda)}\\ &=\frac{(-1)^\beta}{2^\beta\Gamma(\beta+1)}\frac{\Gamma(\lambda+\beta+1)}{\Gamma(\lambda-\beta+1)}\end{aligned} \tag{3.62}$$

when $\beta\in\mathbb{Z}_+$, $-\lambda+\beta$ and $-\lambda$ are not integers.

Theorem 3.19. *Let $h\in\mathcal{B}\cap C^\infty(\mathbb{S}^2)$ be admissible such that there exists $\nu\in(0,\frac{\pi}{2})$ such that $h(\theta^{'},\varphi^{'})=0$ for $|\varphi^{'}|<\nu$ and for $|\varphi^{'}-\pi|<\nu$. Assume that, on subset $\Lambda\neq\emptyset$ of integer pairs (λ_0,m_0) such that $\left|\langle e^{-im_0\varphi^{'}}h,Y_{\lambda_0}^{m_0}\rangle\right|=b_{\lambda_0,m_0}$ for some $b_{\lambda_0,m_0}>0$. Then for any $f\in\bigoplus_{m\leqslant M}\mathfrak{M}_m$ and $\alpha=\frac{1}{2}$, there exists $0<A\leqslant B<\infty$ such that*

$$A\|f\|^2\leqslant\int_0^\infty\frac{da}{a^3}\int_{SO(3)}|\langle f,\psi_{\sigma,a}\rangle|^2d\sigma\leqslant B\|f\|^2 \tag{3.63}$$

where

$$\psi(\theta^{'},\varphi^{'})=h(\theta^{'},\varphi^{'})\sqrt{\frac{\theta'}{\sin\theta'}}(1+J^2)\cos^2\frac{\theta^{'}}{2} \tag{3.64}$$

In particular we can take

$$A=\sum_{(\lambda_0,m_0)\in\Lambda}\gamma_{\lambda_0,m_0}\left(\prod_{s=1}^{m_0}\frac{\lambda_0}{\lambda_0+s+1}\right)^2\frac{\ln(\lambda_0+1/2)}{\ln(\lambda_0+1/2-\epsilon_{\lambda_0,m_0})} \tag{3.65}$$

with ϵ_{λ_0,m_0} solely depending on b_{λ_0,m_0} and $\gamma_{\lambda_0,m_0}=\frac{b^2}{(2\lambda_0+2)^{2m_0+2}}\frac{(\lambda_0+m_0-1)!}{(\lambda_0-m_0+1)!}$.

Proof. Suppose that $\left|\langle e^{-im_0\varphi'}h, Y_{\lambda_0}^{m_0}\rangle\right| = b$ for some integer $\lambda_0 \in \Lambda$ and $b > 0$. Take any large $l \in \mathbb{N}$ there exists a_0 such that $\lambda_0 + \frac{1}{2} = \sqrt{a_0}(l+\frac{1}{2})$. Let $\lambda(a)+\frac{1}{2} = \sqrt{a}(l+\frac{1}{2})$, then there exists $\delta \in (0, a_0/2)$ such that for $a \in (a_0-\delta, a_0+\delta)$ there is $|\lambda - \lambda_0| < \epsilon$ for some fixed $\epsilon \in (0, 1/4)$ and

$$\left|\langle e^{-im_0\varphi'}h, Y_{\lambda}^{m_0}\rangle\right| > \frac{b}{2}$$

For instance we can choose $\delta = \frac{2\epsilon(\lambda_0+1/2)-\epsilon^2}{(l+1/2)^2}$. Notice that

$$\frac{a^{-1/2}(\lambda+\frac{1}{2}) - \frac{1}{2} + k}{\lambda+k} = a^{-1/2}\left(\frac{\lambda+1/2}{\lambda+k}\right) + O(1)$$

for any $1 \leqslant k \leqslant 2m-1$ and that

$$\begin{aligned}
P_l^{m_0}(\cos\theta) = & \frac{\left(a^{-1/2}(\lambda+\frac{1}{2})+\frac{1}{2}\right)\left(a^{-1/2}(\lambda+\frac{1}{2})+\frac{3}{2}\right)\cdots\left(a^{-1/2}(\lambda+\frac{1}{2})+m_0-\frac{1}{2}\right)}{\left(a^{-1/2}(\lambda+\frac{1}{2})\right)^{m_0}} \\
& \frac{\Gamma\left((a^{-1/2}(\lambda+\frac{1}{2})+1\right)}{\left(a^{-1/2}(\lambda+\frac{1}{2})-m_0\right)!}\left(\frac{\theta\sin\theta'}{\theta'\sin\theta}\right)^{\frac{1}{2}}\left(\frac{\theta'}{\sin\theta'}\right)^{\frac{1}{2}} J_{m_0}\left((\lambda+\frac{1}{2})\theta'\right) \\
& + O(a^{m_0/2+1}) \\
& = q(a)\left(\frac{\sin\theta'}{\theta'}\right)^{\frac{1}{2}} P_\lambda^{m_0}(\cos\theta') + O(a^{m_0/2+1})
\end{aligned} \tag{3.66}$$

where $q(a) = a^{-m_0/2}\prod_{s=0}^{m_0-1}\frac{\lambda+1/2}{\lambda+s+1} + O(a^{-\frac{m_0-1}{2}})$.

From the previous analyzing we know that for ψ admissible $\widehat{(\psi_a)}_{l0}$ vanishes, hence we can without loss of generality assume that m_0 is positive. Since $\psi(\theta', \varphi') = 0$ for $|\varphi'| < \nu$ and for $|\varphi' - \pi| < \nu$, along with $\alpha = \frac{1}{2}$, there is

$$\begin{aligned}
I_{l,m_0} &= \frac{1}{2l+1}\int_0^\infty \left|\widehat{(\psi_a)}_{l,m_0}\right|^2 \frac{da}{a^3} \\
&\geqslant \frac{4|c_{l,m_0}|^2}{(2l+1)}\int_{a_0-\delta}^{a_0+\delta} |c_{\lambda,m_0}|^{-2}\left|\langle e^{-im_0\varphi'}h, Y_\lambda^{m_0}\rangle\right|^2 q^2(a)a^{-3/2}da \\
&\geqslant \frac{b^2(l-m_0)!}{4\pi(l+m_0)!}\int_{a_0-\delta}^{a_0+\delta} |c_{\lambda,m_0}|^{-2}\left(\prod_{s=0}^{m_0-1}\frac{\lambda+1/2}{\lambda+s+1}\right)^2 a^{-m_0-\frac{3}{2}}da \\
&\geqslant \frac{b^2(l-m_0)!}{(l+m_0)!}\frac{(\lambda_0+m_0-1)!}{(2\lambda_0+1)(\lambda_0-m_0+1)!}\int_{a_0-\delta}^{a_0+\delta}\left(\prod_{s=0}^{m_0-1}\frac{\lambda+1/2}{\lambda+s+1}\right)^2 a^{-m_0-\frac{3}{2}}da
\end{aligned}$$

Using the estimations

$$(2l+1)^{-2m_0} \leqslant (2l^2)^{-m_0} \leqslant \frac{(l-m_0)!}{(l+m_0)!} \tag{3.67}$$

and

$$\frac{(l-m_0)!}{(l+m_0)!} \leqslant l^{-2m_0} \leqslant 2^{m_0} \left(l+\frac{1}{2}\right)^{-2m_0} \quad \text{for} \quad l \geqslant m_0^2 \tag{3.68}$$

it follows that

$$I_{l,m_0} \geqslant \frac{b^2}{(2\lambda_0+2)^{2m_0+2}} \frac{(\lambda_0+m_0-1)!}{(\lambda_0-m_0+1)!} \left(\prod_{s=1}^{m_0} \frac{\lambda_0}{\lambda_0+s+1}\right)^2 \ln \frac{a_0+\delta}{a_0-\delta}$$

Furthermore, by the choice of δ, we have

$$\ln\left(\frac{a_0+\delta}{a_0-\delta}\right) \geqslant 2\left[\ln(\lambda_0+1/2) - \ln(\lambda_0+1/2-\epsilon)\right]$$

independent of l. Thus we have arrived at (3.65) when we consider all those integer pairs (λ_0, m_0).

For an arbitrary pair $(\lambda_0, m_0) \in \Lambda$, due to our analyzing in Theorem 2.15, the estimation (3.68) and the observation

$$|c_{\lambda,m_0}|^{-2} \leqslant \frac{4\pi(\lambda_0+m_0+1)!}{(2\lambda_0+1/2)(\lambda_0-m_0-1)!} \tag{3.69}$$

we have for $l \geqslant m_0^2$ that

$$\begin{aligned} I_{l,m_0}[0,\alpha_0] &= \frac{1}{2l+1} \int_0^{a_0} \left|\widehat{(\psi_a)}_{l,m_0}\right|^2 \frac{da}{a^3} \\ &\leqslant \frac{8|c_{l,m_0}|^2}{(2l+1)} \int_0^{a_0} |c_{\lambda,m_0}|^{-2} \left|\langle e^{-im_0\varphi'} h, Y_\lambda^{m_0}\rangle\right|^2 q^2(a) a^{-3/2} da \\ &\leqslant 2^{2m_0+2}(l+\frac{1}{2})^{-2m_0-1} \frac{(\lambda_0+m_0+1)!}{(2\lambda_0+\frac{1}{2})(\lambda_0-m_0-1)!} \left(\prod_{s=0}^{m_0-1} \frac{1}{s+1}\right)^2 \times \\ &\int_0^{a_0} (\lambda+1/2)^{2m_0} \left|\langle e^{-im_0\varphi'} h, Y_\lambda^{m_0}\rangle\right|^2 a^{-m_0-3/2} da \end{aligned} \tag{3.70}$$

Let $\lambda+\frac{1}{2}=\bar{\lambda}$, $d\bar{\lambda}=\frac{\bar{\lambda}}{2a}da$, hence

$$\begin{aligned}&\int_0^{a_0}(l+\frac{1}{2})^{-2m_0-1}(\lambda+1/2)^{2m_0}\left|\langle e^{-im_0\varphi'}h,Y_\lambda^{m_0}\rangle\right|^2a^{-m_0-3/2}da\\ \leqslant&\frac{1}{2}\|h\|_1^2\int_0^{\lambda_0+\frac{1}{2}}\|Y_\lambda^{m_0}\|_\infty^2\bar{\lambda}^{2m_0}\bar{\lambda}^{-2m_0}d\bar{\lambda}\\ \leqslant&\frac{1}{2}\|h\|_1^2\sup_{\lambda\in[0,\lambda_0+\frac{1}{2}]}\|Y_\lambda^{m_0}\|_\infty^2(\lambda_0+\frac{1}{2})\end{aligned}$$

Thus we claim that $I_{l,m_0}[0,\alpha_0]$ is bounded. Furthermore, the assumption that $\psi\in C^1(\mathbb{S}^2)$ immediately leads to $\|\mathbb{P}_{\lambda_0}\psi_a\|^2=O(\lambda_0^{-1})$, hence the integral on the domain $(a_0,1)$ is bounded by some constant. Finally the fact that $\|\mathbb{P}_l\psi\|\leqslant\|\psi\|$ suffices to conclude the boundedness on $(1,\infty)$. In sum, the integral $\frac{1}{2\lambda_0+1}\int_0^\infty|\mathbb{P}_{\lambda_0}\psi_a|^2a^{-3}da$ is bounded by some $B<\infty$ independent of λ_0 and it finishes the proof. □

Combining the above result with Proposition 2.11 we arrive at a reproducing formula:

Corollary 3.20. *Given ψ that satisfies the conditions in Theorem 3.19, there is*

$$\int_0^\infty\frac{da}{a^3}\int_{SO(3)}\langle f,\psi_{\sigma,a}^\sharp\rangle\psi_{\sigma,a}^\sharp d\sigma=f$$

for any $f\in L^2(\mathbb{S}^2)$, where $\mathbb{P}_l\psi^\sharp:=\mathbb{P}_l\psi/\sqrt{C_\psi^l}$ for each l.

In practice we cannot do integration on the whole domain, but rather recourse to a discretized summation to approximate the smooth case.

Theorem 3.21. *Under the same assumption of Theorem 3.19, let $\Upsilon_{\delta,\epsilon}$ be the set of sequences $\{a_j\}_{j\geqslant0}$ such that $a_j\to0$, $|a_j-a_{j+1}|\leqslant\delta$, $a_j^{3\alpha-1}/a_{j+1}^2\leqslant C$ and $1\leqslant a_0^2\epsilon$ with δ and ϵ sufficiently small. Then for each $\{a_j\}_j\in\Upsilon_{\delta,\epsilon}$, $\{\sqrt{s_j}\psi_{a_j}\}_{j\in\mathbb{N}}$ forms a semi-frame system for $L^2(\mathbb{S}^2)$, where $s_j=\frac{1}{a_{j+1}^2}-\frac{1}{a_j^2}$.*

Proof. Firstly notice that $\frac{d\cos\theta}{da}=O(a^{2\alpha-1})$ and that

$$\begin{aligned}\frac{d}{d\cos\theta}P_l^m(\cos\theta)=&-m\cos\theta\sin^{m-2}\theta\frac{\Gamma(l+m+1)}{2^m\Gamma(l+1)}P_{l-m}^{m,m}(\cos\theta)\\&+\sin^m\theta\frac{\Gamma(l+m+2)}{2^{m+1}\Gamma(l+1)}P_{l-m-1}^{m+1,m+1}(\cos\theta)\\=&O(a^{\alpha(m-2)})\end{aligned}$$

hence $\frac{d}{da}P_l^m(\cos\theta) = O(a^{\alpha m-1})$. Since $\psi(\theta,\varphi)=0$ for $\varphi<\nu$, there is

$$\frac{d}{da}e^{-im\varphi} = -me^{-i(m+1)\varphi}\left(\frac{i\tan\varphi' + a^{1-\alpha}}{\pm\sqrt{\tan^2\varphi' + a^{2-2\alpha}}}\right)' = O(a^{-\alpha})$$

Besides, a direct calculation gives $\frac{d}{da}(1+J^2) = O(a^{2\alpha-1})$. From (2.66) there is $\left(\widehat{\psi_a}\right)'_{lm} = O(a^{(\alpha-1)/2+\alpha m})$, hence combined with our previous result we arrive at $\left|\left(\widehat{\psi_a}\right)^2_{lm}\right|' = O(a^{\alpha(2m+1)})$. Thus

$$\begin{aligned}
&\frac{1}{2l+1}\left|\frac{1}{2}\sum_{1\leqslant|m|\leqslant l}\sum_{j=0}^{\infty}\left|\left(\widehat{\psi_{a_j}}\right)_{lm}\right|^2\left(\frac{1}{a_{j+1}^2}-\frac{1}{a_j^2}\right) - \sum_{|m|\leqslant l}\int_0^\infty \frac{d}{da^3}\left|\left(\widehat{\psi_a}\right)_{lm}\right|^2\right| \\
&\leqslant \frac{1}{2l+1}\sum_{1\leqslant|m|\leqslant l}\left|\sum_{j=0}^{\infty}\int_{a_{j+1}}^{a_j}\int_a^{a_j}\left|\left(\widehat{\psi_s}\right)^2_{lm}\right|' ds\frac{da}{a^3}\right| + \frac{1}{2l+1}\sum_{1\leqslant|m|\leqslant l}\int_{a_0}^\infty \frac{da}{a^3}\left|\left(\widehat{\psi_a}\right)_{lm}\right|^2 \\
&\leqslant \sum_{j=0}^{\infty}\frac{\left(a_j-a_{j+1}\right)^2 a_j^{3\alpha-1}}{a_{j+1}^2} + \epsilon
\end{aligned}$$

where ϵ is arbitrarily small for sufficiently large a_0. Furthermore, under our assumption that $|a_j - a_{j+1}| \leqslant \delta$ and $\frac{a_j^{3\alpha-1}}{a_{j+1}^2} \leqslant C$, the last line of the above inequality is bounded by $C\delta+\epsilon$, thus the conclusion holds.

□

The next result from [62] provides us one candidate sampling method on the rotation group. I include its simple proof here for completeness.

Lemma 3.22. *If the quadrature formulae*

$$\frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} Y_l^m(\varphi,\theta)\sin\theta d\theta d\varphi = \sum_{i\in I_1} w_i(\mathbb{S}^2)Y_l^m(\varphi_i,\theta_i) \tag{3.71}$$

with $l\leq N$ and

$$\frac{1}{2\pi}\int_0^{2\pi} e^{in\phi}d\phi = \sum_{j\in I_2} w_j(\mathbb{S}^1)e^{in\phi_j} \tag{3.72}$$

with $|n|\leq N$, then all polynomials $f\in\oplus_{l\leq N}\mathcal{H}_l$ can be expressed by

$$\int_{SO(3)} f d\sigma = \sum_{i\in I_1}\sum_{j\in I_2} w_i(\mathbb{S}^2)w_j(\mathbb{S}^1)f(\overline{\varphi}_{1i},\overline{\theta}_i,\overline{\varphi}_{2j}) \tag{3.73}$$

where $(\overline{\varphi}_1, \overline{\theta}, \overline{\varphi}_2)$ represents the Euler angles.

Proof. For $n \neq 0$, by (3.71) and (3.72) we have

$$\begin{aligned}
&\frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} D^l_{m,n}(\bar{\varphi}_1, \bar{\theta}, \bar{\varphi}_2)\sin\bar{\theta}d\bar{\varphi}_1 d\bar{\theta}d\bar{\varphi}_2 \\
&= \frac{1}{4\pi}\int_0^{2\pi}\int_0^{\pi} e^{-im\bar{\varphi}_1} d^l_{m,n}(\bar{\theta})\sin\bar{\theta}d\bar{\theta}d\bar{\varphi}_1 \frac{1}{2\pi}\int_0^{2\pi} e^{-in\bar{\varphi}_2}d\bar{\varphi}_2 \\
&= \sum_{i\in I_1} w_i(\mathbb{S}^2)e^{-im\overline{\varphi}_{1i}} d^l_{m,n}(\overline{\theta}_i)\sum_{j\in I_2} w_j(\mathbb{S}^1)e^{-in\overline{\varphi}_{2j}}
\end{aligned}$$

For $n = 0$, note that $D^l_{m,0}(\bar{\varphi}_1, \bar{\theta}, \bar{\varphi}_2) = (-1)^m\sqrt{\frac{4\pi}{2l+1}}Y_l^{-m}(\bar{\varphi}_1, \bar{\theta})$ and $\sum_j w_j(\mathbb{S}^1) = 1$, hence

$$\begin{aligned}
&\frac{1}{8\pi^2}\int_0^{2\pi}\int_0^{\pi}\int_0^{2\pi} D^l_{m,0}(\bar{\varphi}_1, \bar{\theta}, \bar{\varphi}_2)\sin\bar{\theta}d\bar{\varphi}_1 d\bar{\theta}d\bar{\varphi}_2 \\
&=\sum_i w_i(\mathbb{S}^2)(-1)^m\sqrt{\frac{4\pi}{2l+1}}Y_l^{-m}(\bar{\varphi}_{1i}, \bar{\theta}_i) \\
&= \sum_i\sum_j w_i(\mathbb{S}^2)w_j(\mathbb{S}^1)D^l_{m,0}(\bar{\varphi}_{1i}, \bar{\theta}_i, \bar{\varphi}_{2j})
\end{aligned}$$

Now (3.73) is an immediate result of the expression that

$$f = \sum_l d_l \sum_m\sum_n \langle f, D^l_{m,n}\rangle_{SO(3)} D^l_{m,n} \tag{3.74}$$

□

This lemma enables us to have a fully discretized frame, as a consequence of Theorem 3.19. Let

$$C^l_{\psi,\Upsilon} = \frac{1}{2l+1}\sum_{a_k\in\Upsilon_{\delta,\epsilon}} s_k\|\mathbb{P}_l\psi_{a_k}\|^2 \tag{3.75}$$

Corollary 3.23. *Under the assumptions of Theorem* 3.19 *and Theorem* 3.21*,*

$$\{\psi_{i,j,k} = \sqrt{w_i(\mathbb{S}^2)w_j(\mathbb{S}^1)s_k}\mathcal{R}(\sigma_{i,j})\psi_{a_k} : i \in I_1, j \in I_2, \{a_k\}_{k\geqslant 0} \in \Upsilon_{\delta,\epsilon}\}$$

is a fully discrete frame for $L^2(\mathbb{S}^2)$, where $\sigma_{i,j} = (\varphi_i, \theta_i, \phi_j)$. In particular, let

$$\mathbb{P}_l\psi^\sharp_{i,j,k} := \mathbb{P}_l\psi_{i,j,k}/\sqrt{C^l_{\psi,\Upsilon}} \tag{3.76}$$

then $\{\psi^\sharp_{i,j,k}\}_{i,j,k}$ *is super tight, and there is the reproducing formula:*

$$f = \sum_{i' \in I_1} w_{i'}(\mathbb{S}^2) \sum_{i \in I_1} \sum_{j \in I_2} \sum_{a_k \in \Upsilon} f(\varphi_{i'}, \theta_{i'}) \psi^\sharp_{i,j,k}(\varphi_{i'}, \theta_{i'}) \psi^\sharp_{i,j,k} \tag{3.77}$$

for any $f \in L^2(\mathbb{S}^2)$.

Proof. Similar to Proposition 2.11, we can prove that

$$\sum_{i \in I_1} \sum_{j \in I_2} \sum_{a_k \in \Upsilon_{\delta,\epsilon}} \langle f, \psi_{i,j,k} \rangle_{L^2(\mathbb{S}^2)} \psi_{i,j,k} = \sum_{l} \sum_{|m| \leqslant l} C^l_{\psi,\Upsilon} \hat{f}_{lm} Y^m_l \tag{3.78}$$

Since Theorem 3.21 implies that $C^l_{\psi,\Upsilon} > 0$, there is

$$f = \sum_{i \in I_1} \sum_{j \in I_2} \sum_{a_k \in \Upsilon} \langle f, \psi^\sharp_{i,j,k} \rangle \psi^\sharp_{i,j,k} \tag{3.79}$$

Using (3.71) again leads to the claimed result (3.77).

□

I do not intend to give redundancy analysis of spherical frames in this dissertation, and their numerical simulations are also left to my future work. However, at this point it is appropriate to mention briefly their connection to solving PDEs. Consider an operator equation

$$\mathcal{S}u = g \tag{3.80}$$

where $\mathcal{S}$ can be a differential operator or integral operator from $L^2(\mathbb{S}^2)$ into itself.

Suppose that we have a multi-resolution structure such that $L^2(\mathbb{S}^2) = \bigoplus_{j=0}^{\infty} W_j$ in which each W_j has a finite frame. I have given such a structure for Shannon type spherical wavelets in Example 3.17, though temporarily it has not been done for spherical α-wavelets/shearlets. In this situation we can propose a Galerkin scheme

$$\langle \mathcal{S}u, \nu \rangle = \langle g, \nu \rangle \text{ for } \nu \in V_j \tag{3.81}$$

or

$$\langle \mathcal{S}u, \psi \rangle = \langle g, \psi \rangle \text{ for } \psi \in W_j \tag{3.82}$$

for $j = 0, \cdots, J$, and denote by u^J the solution of this finite element formulation. In particular, under the assumption that $\{\psi^j_i\}_{i \in I^j}$ is a super tight frame for W_j, if $\mathcal{S}$ is invertible, then there are explicite expressions for u^J and u, namely

$$u^J = \sum_{j=0}^{J} \sum_{i \in I^j} \langle g, (\mathcal{S}^{-1})^* \psi^j_i \rangle \psi^j_i \tag{3.83}$$

and

$$u = \sum_{j=0}^{\infty} \sum_{i \in I^j} \langle g, (\mathcal{S}^{-1})^* \psi_i^j \rangle \psi_i^j \tag{3.84}$$

although practically it is often not clear what is the inverse operator hence not recommendable for computation.

Proposition 3.24. *Let $\mathbb{P}_{W_j}$ denote the projection onto W_j, we have the estimation for the error $e^J = u - u^J$ that*

$$\|e^J\|_{L^2} = \frac{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{W_j} \mathcal{S}^{-1} g\|_{L^2}^2}{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{W_j} \mathcal{S}^{-1} g\|_{L^2}}$$

If $\mathcal{S}$ is a unitary operator, then

$$\|e^J\|_{L^2} = \frac{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{\mathcal{S}W_j} g\|_{L^2}^2}{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{\mathcal{S}W_j} g\|_{L^2}}$$

Proof. This is an immediate consequence of the observation that the left-hand side of (3.24) is equal to

$$\begin{aligned}
\sup_{\|\varphi\|=1} \left| \langle e^J, \varphi \rangle \right| &= \sup_{\|\varphi\|=1} \sum_{j=J+1}^{\infty} \sum_{i \in I^j} \left| \langle \mathcal{S}^{-1} g, \psi_i^j \rangle \langle \psi_i^j, \varphi \rangle \right| \\
&= \sup_{\|\varphi\|=1} \sum_{j=J+1}^{\infty} \left| \langle \mathbb{P}_{W_j} \mathcal{S}^{-1} g, \varphi \rangle \right|_{L^2} \\
&= \frac{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{W_j} \mathcal{S}^{-1} g\|_{L^2}^2}{\sum\limits_{j=J+1}^{\infty} \|\mathbb{P}_{W_j} \mathcal{S}^{-1} g\|_{L^2}}
\end{aligned}$$

Similarly when $\mathcal{S}$ is unitary, there is

$$\begin{aligned}
\|e^J\|_{L^2} &= \sup_{\|\varphi\|=1} \sum_{j=J+1}^{\infty} \sum_{i \in I^j} \left| \langle g, \mathcal{S} \psi_i^j \rangle \langle \mathcal{S} \psi_i^j, \mathcal{S} \varphi \rangle \right| \\
&= \sup_{\|\varphi\|=1} \sum_{j=J+1}^{\infty} \left| \langle \mathbb{P}_{\mathcal{S}W_j} g, \mathcal{S} \varphi \rangle \right|_{L^2}
\end{aligned}$$

which is further equal to the right-hand side of (3.24). □

3.3 A product formula on simple surfaces

To avoid confusion with previous sections, throughout this section the symbol Γ is reserved for polyhedra and T denotes triangles on surfaces. For topologists, triangulation of a topological manifold $\mathbb{M}$ means a simplicial complex together with a homeomorphism from its geometric realization to the manifold. It is a deep result in topology that every smooth surface has a triangulation and manifolds of dimension less than four has a piecewise linear triangulation. Nevertheless, I have no intention to delve into topological context or any of these advanced questions, because the purpose in this small section is to establish a formula for a newly introduced inner product on surfaces.

According to the definition from [70], a set of triangles in $\mathbb{R}^2$ is a *triangulation* if
(i) for $i \neq j$, $T_i \cap T_j$ is either empty or a common vertex or a common edge
(ii) the number of boundary edges incident on a boundary vertex is two
(iii) $\bigcup\limits_i T_i$ is simply connected.

Meanwhile a triangulation can be considered as embedding of the graph G and here we additionally assume that
(iv) every vertex has finite degree or G is locally finite.

A more general concept for a graph embedded into an oriented surface without self-intersection is the so called *tessellation*, where the faces are not limited to triangular ones, namely a graph that satisfy
(I) every edge is contained in two faces
(II) every two faces are either disjoint or intersect in one vertex or one edge
(III) every face is homeomorphic to a closed disc

Let $\mathcal{N}(v)$ be the set of neighboring vertices of v, and $|\mathcal{N}(v)| = \deg(v)$ be the degree of v. An edge v_1v_2 is *simple* if $v^{'} \in \mathcal{N}(v_1) \cap \mathcal{N}(v_2)$, then $v^{'}v_1v_2 \in \mathcal{T}$. When every edge is simple, the triangulation is called simple. Obviously for an interior edge, it is simple if and only if $|\mathcal{N}(v_1) \cap \mathcal{N}(v_2)|$=2, where $|\cdot|$ means cardinality.

Square summable maps on the graph form a Hilbert space $\ell^2(\mathcal{V})$ with inner product

$$(f, g)_w = \sum_{v \in \mathcal{V}} f(v)g(v)\deg(v) \tag{3.85}$$

and the weighted combinatoric Laplacian reads

$$\Delta_w g(v) = \frac{1}{\deg(v)} \sum_{v' \in \mathcal{N}(v)} \Big(g(v') - g(v) \Big) = -g(v) + \frac{1}{\deg(v)} \sum_{v' \in \mathcal{N}(v)} g(v') \quad (3.86)$$

which is obviously a self-adjoint operator on the space $\{f : \Delta_w f \in \ell^2(\mathcal{V})\}$. While at the same time due to the Green's formula

$$(-\Delta_w f, g)_w = \frac{1}{2} \sum_v \sum_{v'} df([v, v'])dg([v, v'])\delta_{vv'} \quad (3.87)$$

where $df([v, v']) = f(v') - f(v)$ and $\delta_{vv'} = 1$ if v and v' are neighboring, zero otherwise, it follows that $-\Delta_w$ is positive. Thus its spectrum $\sigma(-\Delta_w) = \{z \in \mathbb{C} : zI + \Delta_w$ has no bounded inverse$\}$ lies on the positive part of the real line. In fact, Theorem 4.18 gives bounds estimation for $\sigma(-\Delta_w)\backslash\{0\}$, particularly the lower bound for the first minimal eigenvalue that is larger than zero.

By using radial projection method, once given a polyhedron with triangular facades $\mathcal{T} = \{T_1, \cdots, T_N\}$, one get a natural triangulation $p(\mathcal{T})$ on the sphere. A norm adapted to the triangulation, as the square root of a newly introduced inner product in $L^2(\partial\Gamma)$ is defined as

$$\langle f, g\rangle_{\partial\Gamma} = \sum_{T \in \mathcal{T}} \frac{1}{|T|} \int_T f(x)g(x)dx \quad (3.88)$$

where we denote by $\partial\Gamma$ the surface of the polyhedron Γ, to distinguish this norm from the $L^2(\mathbb{M})$ norm on the Riemannian manifold.

The pull-back of the norm (3.88) by p then gives a norm on the sphere. Furthermore, a continuous function f piecewise-linear on triangle in $\mathcal{T}$ is uniquely determined by the values $f(\mathcal{V})$, where $\mathcal{V}$ is the set of vertices in the triangulation. Denote by the space of these functions by $\mathbb{P}^1$, we generalize a useful result in [28] and give a different but simpler proof here

Proposition 3.25. *Let $\mathcal{T}$ be a simple triangulation of an oriented surface without boundary, f and g be two elements in $\mathbb{P}^1(\mathcal{T})$, and denote by $\mathcal{V}_i = \{v_k^i\}_{k=1,2,3}$ the set of vertices of triangle T_i. Then*

$$\langle f, g\rangle_{\partial\Gamma} = \frac{1}{24} \left[2(f, g)_w + \sum_{v \in \mathcal{V}} \sum_{v' \in \mathcal{N}(v)} \Big(f(v)g(v') + g(v)f(v') \Big) \right] \quad (3.89)$$

Proof. Let ϕ_s and ϕ_t be two arbitrary nodal functions at vertices s and t respectively. Firstly observe that, if s and t are contained in the same triangle T_i, then

$$\begin{aligned}
&\frac{1}{2|T_i|}\int_{T_i}\phi_s(x)\phi_t(x)dx \\
&= \int_0^1\int_0^{1-\lambda_2}(\phi_s\cdot\phi_t)(\lambda_1 v_1^i+\lambda_2 v_2^i+(1-\lambda_1-\lambda_2)v_3^i)d\lambda_1 d\lambda_2 \\
&= \begin{cases} \frac{1}{24} & s\sim t \\ \frac{1}{12} & s=t \end{cases}
\end{aligned}$$

where $s\sim t$ means that they are neighboring vertices. Consequently,

$$\begin{aligned}
I_{s,t} = \sum_{T_i\in\mathcal{T}}\frac{1}{2|T_i|}\int_{T_i}\phi_s(x)\phi_t(x)dx &= \sum_{i\in\mathcal{N}(s)\cap\mathcal{N}(t)}\frac{1}{2|T_i|}\int_{T_i}\phi_s(x)\phi_t(x)dx \\
&= \begin{cases} 1/12 & s\sim t \\ \deg(s)/12 & s=t \end{cases}
\end{aligned}$$

In the situation that s and t are not contained in any triangle, clearly there is $I_{s,t}=0$.

Since the nodal functions form a basis for $\mathbb{P}^1$ and a piecewise linear function restricted to T_i is the linear combination of $\{\phi_v\}_{v\in\mathcal{V}_i}$, we arrive at the result that

$$\begin{aligned}
&\sum_{T_i\in\mathcal{T}}\frac{1}{2|T_i|}\int_{T_i}f(x)g(x)dx \\
&= \sum_{T_i\in\mathcal{T}}\frac{1}{2|T_i|}\int_{T_i}\sum_{v\in\mathcal{V}^i}f(v)\phi_v(x)\cdot\sum_{v\in\mathcal{V}_i}g(v)\phi_v(x)dx \\
&= \frac{1}{12}\sum_{v\in\mathcal{V}}\deg(v)f(v)g(v)+\frac{1}{24}\sum_{v\in\mathcal{V}}\sum_{v'\in\mathcal{N}(v)}\Big(f(v)g(v')+g(v)f(v')\Big)
\end{aligned}$$

$\square$

Note that $\|f\|^2_{L^2(\mathbb{M})}=\sum_{T_i}\int_{\mathbb{M}}f^2(x)dx\leqslant\max_i|T_i|\langle f,f\rangle_{\partial\Gamma}$. Thus when f has normalized $L^2(\mathbb{M})$ norm and the triangulation becomes finer and finer such that $\max_i|T_i|$ turns smaller and smaller, $\langle f,f\rangle_{\partial\Gamma}$ becomes unbounded. In other words, they are not comparable on the same scale. However, Proposition 3.25 does allow us to compare $\langle f,f\rangle_{\partial\Gamma}$ with $(f,f)_w$.

Corollary 3.26. *Under the assumption that each vertex in the triangulation has degree no bigger than p for some $p \geqslant 2$, there is*

$$1 + \frac{2}{p} \leqslant \frac{12\langle f, f\rangle_{\partial\Gamma}}{(f, f)_w} \leqslant 2 \tag{3.90}$$

for any $f \in \mathbb{P}^1(\mathcal{T}) \subset L^2(\partial\Gamma)$ not identically zero on $\mathcal{V}$.

Proof. In the case that $f = g$, (3.89) reduces to

$$\langle f, f\rangle_{\partial\Gamma} = \frac{1}{12}\left[(f, f)_w + \sum_v \sum_{v'} f(v) f(v^{'})\delta_{v,v'}\right] \tag{3.91}$$

Thus on the right side

$$\begin{aligned}
\langle f, f\rangle_{\partial\Gamma} &\leqslant \frac{1}{12}\left[(f, f)_w + \sum_v \sum_{v'} f(v) f(v^{'})\delta_{v,v'}\right] \\
&\leqslant \frac{1}{12}\left[(f, f)_w + \frac{1}{2}\sum_v \sum_{v'} (f(v)^2 + f(v^{'})^2)\delta_{v,v'}\right] \\
&= \frac{1}{6}(f, f)_w
\end{aligned}$$

while the left side of (3.90) is obvious. □

While in differential geometry there are equivalent descriptions of one mathematical concept, in the discrete world different definitions motivated by the various smooth properties do not usually lead to the same thing. For instance in [8] a discrete Laplace operator is constructed through the discretization of mean curvature, while another discrete Laplace operator can be formulated based on a discrete analogue of Green's theorem(see [4] for instance),

$$\sum_i | \nabla g|_{T_i} |^2 \cdot |T_i| = -\sum_j g(v_j)\Delta_{\mathcal{T}} g(v_j) w(v_j) \tag{3.92}$$

where $w(v_j)$ is some area weight at vertices v_j.

On an oriented surface which has a consistent triangulation, the discrete gradient of a function g on $\mathbb{M}$ can be defined at each triangle T_i as the solution of

$$\langle \nabla g|_{T_i}, v_k^i - c^i\rangle = g(v_k^i) - g(c^i) \quad i = 1, 2, 3 \tag{3.93}$$

where $c^i = \frac{1}{3}\sum_{k=1}^{3}(v_k^i)$ and the inner product is in the Euclidean sense. From this definition it is clear that $\langle \nabla g|_{T_i}, v_j^i - v_k^i\rangle = g(v_j^i) - g(v_k^i)$. Suppose $e_1^i = v_3^i - v_2^i$, $e_2^i = v_1^i - v_3^i$, $e_3^i = v_2^i - v_1^i$ and $\nabla g|_{T_i} = c_1 e_1^i + c_2 e_2^i$, then

$$G\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} \langle \nabla g|_{T_i}, e_1^i\rangle \\ \langle \nabla g|_{T_i}, e_2^i\rangle \end{pmatrix} = \begin{pmatrix} g(v_3^i) - g(v_2^i) \\ g(v_1^i) - g(v_3^i) \end{pmatrix} \tag{3.94}$$

where

$$G = \begin{pmatrix} \langle e_1^i, e_1^i\rangle & \langle e_1^i, e_2^i\rangle \\ \langle e_1^i, e_2^i\rangle & \langle e_2^i, e_2^i\rangle \end{pmatrix}$$

Thus, if we omit the index i without bringing confusion

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{1}{\det G}\cdot\begin{pmatrix} \langle e_2, e_2\rangle & -\langle e_1, e_2\rangle \\ -\langle e_1, e_2\rangle & \langle e_1, e_1\rangle \end{pmatrix}\begin{pmatrix} g(v_3) - g(v_2) \\ g(v_1) - g(v_3) \end{pmatrix}$$

and

$$\begin{aligned}(e_1, e_2)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = {} & \frac{1}{\det G}[-(\langle e_2, e_2\rangle e_1 - \langle e_1, e_2\rangle e_2)\, g(v_2) \\ & + (\langle e_1, e_1\rangle e_2 - \langle e_1, e_2\rangle e_1)\, g(v_1) \\ & - (\langle e_3, e_2\rangle e_1 - \langle e_3, e_1\rangle e_2)\, g(v_3)]\end{aligned}$$

Note that $[e_1, e_2] = e_1 e_2^T - e_2 e_1^T = e_3^T e_1^T - e_1^T e_3^T = e_2^T e_3^T - e_3^T e_2^T$, hence

$$\nabla g|_{T_i} = (e_1, e_2)\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{-1}{4|T_i|^2}[e_1, e_2]\cdot(g(v_1)e_1 + g(v_2)e_2 + g(v_3)e_3) \tag{3.95}$$

since $\det G = |e_1|^2|e_2|^2\sin^2\theta_3 = 4|T_i|^2$.

There is

$$\begin{aligned} & |\nabla g|_{T_i}|^2 \\ & = (c_1 e_1 + c_2 e_2)^T(c_1 e_1 + c_2 e_2) \\ & = (c_1\ c_2)G(c_1\ c_2)^T \\ & = \frac{1}{4|T_i|^2}\begin{pmatrix} g(v_3) - g(v_2) \\ g(v_1) - g(v_3) \end{pmatrix}^T\begin{pmatrix} \langle e_2, e_2\rangle & -\langle e_1, e_2\rangle \\ -\langle e_1, e_2\rangle & \langle e_1, e_1\rangle \end{pmatrix}\begin{pmatrix} g(v_3) - g(v_2) \\ g(v_1) - g(v_3) \end{pmatrix} \\ & = \frac{1}{4|T_i|}\mathbf{g}_i^T\mathbf{E}^i\mathbf{g}_i \end{aligned} \tag{3.96}$$

where $\mathbf{g}_i = (g(v_1), g(v_2), g(v_3))^T$ and $\mathbf{E}^i$ is 3×3 matrix with elements $\mathbf{E}_{j,k}^i = \langle e_j, e_k\rangle/|T_i|$.

Therefore,

$$\sum_{T_i \in \mathcal{T}} |\nabla g|_{T_i}|^2 \cdot |T_i| = \frac{1}{4} \sum_{v \in \mathcal{V}} g(v) \sum_{T_i \ni v} \left[|e_v|^2 g(v) + \langle e_v, e_{v'} \rangle g(v^{'}) + \langle e_v, e_{v''} \rangle g(v^{''}) \right] / |T_i|$$

which implies that we shall define the discrete Laplace-Beltrami operator at v as

$$\Delta_{\mathcal{T}} g|_v = -\frac{1}{4w_v} \sum_{T_i \ni v} \frac{1}{|T_i|} \left[|e_v^i|^2 g(v) + \langle e_v^i, e_{v'}^i \rangle g(v^{'}) + \langle e_v^i, e_{v''}^i \rangle g(v^{''}) \right] \tag{3.97}$$

where $v^{'}$ and $v^{''}$ are the other two vertices in the triangle T_i. It is clear that when g is a constant function the above equality gives $\Delta_{\mathcal{T}} g \equiv 0$.

As a small step forward and a wonderful ending for this chapter, I give the eigenvalues estimation for $\Delta_{\mathcal{T}}$. Without loss of generality, assume that $w(v) = 1$ for all vertices. Let us also assume that the restriction

$$\frac{h}{r} \leqslant C_{\mathcal{T}} \tag{3.98}$$

applies to all the triangles uniformly, where h is the mesh size and r is the radius of the largest interior circle inside a triangle.

Definition 3.27. Let $\mathbb{U}$ be the subspace of $L^2(\mathbb{M})$ consisting of functions g such that on any T_i, $\left(g|_{T_i}(v_{i0}), g|_{T_i}(v_{i1}), g|_{T_i}(v_{i2})\right)^T \notin \ker \mathbf{E}^i$.

Theorem 3.28. *On a triangular mesh satisfying* (3.98), *there are*

$$\lambda_{min}(\mathcal{T}) = \inf_{g \in \mathbb{U}} \frac{-\sum_j g(v_j) \Delta_{\mathcal{T}} g(v_j)}{\sum_j g^2(v_j)} \geqslant \frac{\bar{p}}{4} \bar{\kappa}(C_{\mathcal{T}})$$

$$\lambda_{max}(\mathcal{T}) = \sup_{g \in \mathbb{U}} \frac{-\sum_j g(v_j) \Delta_{\mathcal{T}} g(v_j)}{\sum_j g^2(v_j)} \leqslant \frac{p}{4} \kappa(C_{\mathcal{T}})$$

where $\bar{\kappa}(C_{\mathcal{T}}) = \sqrt{3} C_{\mathcal{T}} - \sqrt{3C_{\mathcal{T}}^2 - 12}$, $\kappa(C_{\mathcal{T}}) = \sqrt{3} C_{\mathcal{T}} + \sqrt{3C_{\mathcal{T}}^2 - 12}$, *$p$ is the maximum degree and $\bar{p}$ is the minimum degree of the vertices.*

Proof. Suppose the eigenvalues of $\mathbf{E}^i$ are reordered so that $0 \leqslant \lambda_0 \leqslant \lambda_1 \leqslant \lambda_2$, based on the fact that $\mathbf{E}^i$ is positive semi-definite. Let $I = \{0, 1, 2\}$ be the index set and denote by θ_j ($j \in I$) the interior angles of the triangle corresponding to the edge e_j respectively. From the observation that

$$\begin{aligned}\prod_{j=0}^{2} \lambda_j &= \det \mathbf{E}^i \\ &= \frac{1}{|T_i|^3}\left(\prod_{j=0}^{2} |e_j|^2 - \prod_{j=0}^{2} |e_j|^2 \, |\langle e_k, e_l\rangle|^2_{k\neq l\in I\setminus\{j\}} + 2\langle e_0, e_1\rangle\langle e_1, e_2\rangle\langle e_2, e_0\rangle\right) \\ &= \frac{|e_0|^2|e_1|^2|e_2|^2}{|T_i|^3}\left(1-\cos^2\theta_0-\cos^2\theta_1-\cos^2\theta_2-2\cos\theta_0\cos\theta_1\cos\theta_2\right) \\ &= 0\end{aligned} \tag{3.99}$$

we see $\lambda_0 = 0$.

Besides, it holds that

$$\sum_{j=1}^{2} \lambda_j = \mathrm{Tr}\mathbf{E}^i = \frac{\sum\limits_{j=0}^{2} |e_j|^2}{|T_i|^2} = \sum_{j=0}^{2} \frac{2}{\sin\theta_j} := S_i \tag{3.100}$$

and that

$$\lambda_1\lambda_2 = \frac{1}{|T_i|^2}\left(|e_0|^2|e_1|^2\sin^2\theta_2 + |e_1|^2|e_2|^2\sin^2\theta_0 + |e_2|^2|e_0|^2\sin^2\theta_1\right) = 12 \tag{3.101}$$

Equations (3.100) and (3.101) together give us the eigenvalues

$$\begin{cases} \lambda_1 = \frac{S_i - \sqrt{S_i^2-48}}{2} \\ \lambda_2 = \frac{S_i + \sqrt{S_i^2-48}}{2} \end{cases} \tag{3.102}$$

The the assumption (3.98) implies $\frac{1}{\sin\frac{\theta_j}{2}} \leqslant C_{\mathcal{T}}$ for any j, hence

$$\begin{cases} \frac{2}{\sin\theta_j} \leqslant \frac{C_{\mathcal{T}}}{\cos\frac{\theta_j}{2}} \leqslant \frac{2\sqrt{3}C_{\mathcal{T}}}{3} & \theta_j \in (0, \frac{\pi}{3}] \cup [\frac{2\pi}{3}, \pi) \\ \frac{2}{\sin\theta_j} \leqslant \frac{4\sqrt{3}}{3} & \theta_j \in (\frac{\pi}{3}, \frac{2\pi}{3}) \end{cases} \tag{3.103}$$

from which we deduce that

$$\lambda_1 \leqslant 2\sqrt{3} \leqslant \lambda_2 \tag{3.104}$$

Meanwhile it is easy to verify that S_i achieves minimum if and only if $\theta_j = \frac{\pi}{3}$ $(j \in I)$. Thus it follows that

$$4\sqrt{3} \leqslant S_i \leqslant 2\sqrt{3}C_{\mathcal{T}} \tag{3.105}$$

and

$$\begin{cases} \lambda_1 \geqslant \bar{\kappa}(C_{\mathcal{T}}) \\ \lambda_2 \leqslant \kappa(C_{\mathcal{T}}) \end{cases} \tag{3.106}$$

hold for any triangle in the triangulation.

Thus for any $g \in \mathbb{U}\backslash\{0\}$,

$$\frac{-\sum\limits_j g(v_j)\Delta_{\mathcal{T}} g(v_j)}{\sum\limits_j g^2(v_j)} \geqslant \frac{\bar{p} \sum\limits_{T_i \in \mathcal{T}} |\left. \nabla g \right|_{T_i}|^2 \cdot |T_i|}{\sum\limits_{T_i} \sum\limits_{v \in T_i} g^2(v)} \tag{3.107}$$

as a consequence of (3.96). The inequality (3.107) holds because $\sum\limits_{T_i \in \mathcal{T}} |\left. \nabla g \right|_{T_i}|^2 \cdot |T_i| \geqslant \frac{1}{4}\bar{\kappa}(C_{\mathcal{T}}) \cdot \sum\limits_{v \in T_i} g^2(v)$ for each T_i. Similarly we see that

$$\frac{-\sum\limits_j g(v_j)\Delta_{\mathcal{T}} g(v_j)}{\sum\limits_j g^2(v_j)} \leqslant \frac{p \sum\limits_{T_i \in \mathcal{T}} |\left. \nabla g \right|_{T_i}|^2 \cdot |T_i|}{\sum\limits_{T_i} \sum\limits_{v \in T_i} g^2(v)} \tag{3.108}$$

which gives us the estimation (3.28).

□

Remark 3.29. $\mathbb{U}$ automatically excludes piecewise constant functions, which are eigenfunctions of $\Delta_{\mathcal{T}}$ corresponding to eigenvalue 0. Furthermore, $\lambda_{\min}(\mathcal{T})$ can be viewed as the minimum eigenvalue besides zero adapted to a suitable triangulation, it is strictly positive and has lower bound depending on the value of $C_{\mathcal{T}}$ according to (3.28).

Chapter 4

Supporting topics

4.1 Orthogonal polynomials: a differential equation point of view

Sturm-Liouville theory has a long history dating back to the beginning of 19th century, with thousands of papers and articles published under this topic. Here we only give an elegant and short introduction to this profound theory, so as to exhibit how various kinds of orthogonal polynomials can be derived from differential equations. For recent development and ongoing research in this area we refer to monographs [3][96]; for a more detailed introduction we refer to [98].

Consider linear second-order ODE on I(interval, half line, real line etc.)

$$(\tilde{L}y)(x) := a_0(x)y^{''}(x) + a_1(x)y^{'}(x) + a_2(x)y(x) = f(x) \tag{4.1}$$

where $a_0 > 0$ a.e. on I, $\frac{1}{a_0}, a_1 \in L_{loc}(I)$. When necessary we can assume without loss of generality that $a_0 = 1$ and (4.1) is simplified to be

$$L_0 y := y^{''} + a_1 y^{'} + a_2 y = f \tag{4.2}$$

otherwise equation (4.1) is called singular and the zero is called singular point.

Let $x_0,\ a,\ b$ be arbitrary points in I. For any given values c_1 and c_2, the general conditions imposed on (4.1) is

$$\begin{cases} \alpha_1 y(a) + \alpha_2 y^{'}(a) + \alpha_3 y(b) + \alpha_4 y^{'}(b) = c_1 \\ \beta_1 y(b) + \beta_2 y^{'}(b) + \beta_3 y(a) + \beta_4 y^{'}(a) = c_2 \end{cases} \tag{4.3}$$

with $\sum_{i=1}^{4}|\alpha_i| > 0$ and $\sum_{i=1}^{4}|\beta_i| > 0$. Especially, when $\alpha_3 = \alpha_4 = \beta_3 = \beta_4 = 0$ and $a,\ b$ are endpoints of I, we have the *separated boundary conditions*

$$\begin{cases} \alpha_1 y(a) + \alpha_2 y^{'}(a) = c_1 \\ \beta_1 y(b) + \beta_2 y^{'}(b) = c_2 \end{cases} \tag{4.4}$$

when $a = b = x_0$, we have the *initial conditions*

$$y(x_0) = c_1,\ y^{'}(x_0) = c_2 \tag{4.5}$$

and $a_0 > 0$ a.e. on I, $\frac{1}{a_0}, a_1 \in L_{loc}(I)$ are in fact the minimal conditions required to have a unique solution; the special case

$$y(a) = y(b),\ y^{'}(a) = y^{'}(b) \tag{4.6}$$

is called *periodic boundary condition*. Unlike the initial value problems, separated boundary value problems (4.4) do not always have solutions.

Assume $a_0 = 1$ on I. If the coefficients $a_1(x)$ and $a_2(x)$ are both analytic at some point x_0, then we obviously have an (locally) analytic solution of the equation (4.2) of the form $\sum_{n=0}^{\infty} c_n(x - x_0)^n$.

The space of solutions of (4.2) in $L^2(I) \cap C^2(I)$ is closed under linear combination. Integral by parts shows that $\tilde{L} : L^2(I) \cap C^2(I) \mapsto L^2(I)$ satisfies

$$\begin{aligned} \langle \tilde{L}f, g\rangle &= \langle f, (\overline{a_0}g)^{''} - (\overline{a_1}g)^{'} + \overline{a_2}g\rangle + [a_0(f^{'}\overline{g} - f\overline{g}^{'}) + (a_1 - a_0^{'})f\overline{g}]|_a^b \\ &= \langle f, \tilde{L}^{*}g\rangle + [a_0(f^{'}\overline{g} - f\overline{g}^{'}) + (a_1 - a_0^{'})f\overline{g}]|_a^b \end{aligned} \tag{4.7}$$

Then $\tilde{L} = \tilde{L}^{*} \equiv (\overline{a_0})\frac{d^2}{dx^2} + (2\overline{a_0}^{'} - \overline{a_1})\frac{d}{dx} + (\overline{a_0}^{''} - \overline{a_1}^{'} + \overline{a_2})$
$\Leftrightarrow$

$$\overline{a_0} = a_0,\ 2\overline{a_0}^{'} - \overline{a_1} = a_1,\ \overline{a_0}^{''} - \overline{a_1}^{'} + \overline{a_2} = a_2 \tag{4.8}$$

$\Leftrightarrow$

$$a_0,\ a_1 \text{ and } a_2 \text{ are real and } a_1 = a_0^{'} \tag{4.9}$$

In this case, $\tilde{L}$ is called formally self-adjoint and

$$\tilde{L} = \frac{d}{dx}(a_0(x)\frac{d}{dx}) + a_2(x) \tag{4.10}$$

The equation $\tilde{L}y = 0$ with $\tilde{L}$ formally self-adjoint was firstly studied by Charles Sturm and Joseph Liouville in the 1830's, and there all the coefficients and

solutions are of two variables x and r. Under the additional assumption that
1. a_0 is strictly positive and decreasing in r
2. a_2 is increasing function of r
3. $\frac{a_2(a,r)}{y(a,r)}\frac{\partial y}{\partial x}(a,r) = h(r)$ where y solves $Ly = 0$ and h is a given decreasing function of r, Sturm claims that $\frac{a_2(a,r)}{y(a,r)}\frac{\partial y}{\partial x}(a,r) = h(r)$ is decreasing in r for any $x \in I$.

By multiplying a positive weight function w to a general $\tilde{L} = a_0(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_2(x)$, we can make $w\tilde{L}$ be formally self-adjoint. In fact, by (4.9) wL is formally self-adjoint if $w^{'}a_0 + wa_0^{'} = wa_1$, or equivalently

$$|w(x)| = \frac{|a_0(a)w(a)|}{|a_0(x)|}\exp\left(\int_a^x \frac{a_1(t)}{a_0(t)}dt\right) \tag{4.11}$$

Definition 4.1. (Sturm-Liouville Eigenvalue Problem)

Let L be formally self-adjoint operator satisfying (4.10). Solve the eigenvalue equation

$$-Ly = \lambda w(x) y \tag{4.12}$$

on $L^2_w(I)$ subject to the separated homogeneous boundary conditions

$$\begin{cases} \alpha_1 y(a) + \alpha_2 y^{'}(a) = 0 \\ \beta_1 y(b) + \beta_2 y^{'}(b) = 0 \end{cases} \tag{4.13}$$

How about the linear dependence of two solutions of equation (4.12)? If two solutions y_1 and y_2 are linearly dependent, then the Wronskian

$$W(y_1, y_2)(x) = \begin{vmatrix} y_1(x) & y_2(x) \\ y_1^{'}(x) & y_2^{'}(x) \end{vmatrix} \tag{4.14}$$

is identically zero. Conversely, $a_0(y_1y_2^{''} - y_2y_1^{''}) + a_0^{'}(y_1y_2^{'} - y_2y_1^{'}) = 0$ gives us the identity

$$W^{'}(x) + \frac{a_0^{'}(x)}{a_0(x)}W(x) = 0 \tag{4.15}$$

Thus $W(y_1, y_2)(x) = c \cdot \exp(-\int_a^x \frac{a_0^{'}(t)}{a_0(t)}dt)$ for any $x \in I$. $W(y_1, y_2)(x) = 0$ for some $x \Rightarrow c = 0 \Rightarrow W(y_1, y_2)(x) \equiv 0$. Hence we have arrived at the following conclusion.

Lemma 4.2. *Solutions y_1 and y_2 of (4.12) are linearly independent if and only if $W(y_1, y_2)(x) \neq 0$ on I.*

Now according to standard existence theorem of ODE, there exist unique solutions y_1 and y_2 of the homogeneous equation $Ly = 0$ such that $y_1(a) = \alpha_2$, $y_1^{'}(a) = -\alpha_1$; $y_2(b) = \beta_2$, $y_2^{'}(b) = -\beta_1$. Thus

$$\begin{cases} Ly_1 = 0 \\ \alpha_1 y_1(a) + \alpha_2 y_1^{'}(a) = 0 \end{cases} \tag{4.16}$$

and

$$\begin{cases} Ly_2 = 0 \\ \beta_1 y_2(b) + \beta_2 y_2^{'}(b) = 0 \end{cases} \tag{4.17}$$

In order to construct Green's function we need y_1 and y_2 to be linearly independent(y_1 and y_2 are independent iff $\alpha_1 y_2(a) + \alpha_2 y_2^{'}(a) \neq 0$). This is guaranteed under the assumption that $0 \notin spec(L)$, since otherwise v_1 is a multiple of v_2, and v_1 solves both (4.16) and (4.17), namely v_1 solves the SL eigenvalue problem with $\lambda = 0$, a contradiction. Furthermore, as we shall see soon, $0 \notin spec(L)$ is a reasonable assumption.

Definition 4.3. Green's function is defined as

$$G(t,s) = \begin{cases} \frac{y_1(s)y_2(t)}{a_0(s)W(y_1,y_2)(s)} & \text{if} \quad a \leqslant s \leqslant t \leqslant b \\ \frac{y_1(t)y_2(s)}{a_0(s)W(y_1,y_2)(s)} & \text{if} \quad a \leqslant t \leqslant s \leqslant b \end{cases} \tag{4.18}$$

using the fact that $a_0(s) \neq 0$ and $W(y_1, y_2)(s) \neq 0$ on I.

Since

$$[a_0 W(y_1, y_2)]^{'} = y_1 L y_2 - y_2 L y_1 = 0 \tag{4.19}$$

the denominator of Green's function $a_0 W(y_1, y_2)$ turns out to be a nonzero constant c. Thus the following properties of Green's function are satisfied:
(1)G is symmetric
(2)G is continuous on $I \times I$, and belongs to C^2 except for the line $t = s$.

$$\begin{aligned} \lim_{\delta \to 0^+} \frac{\partial G}{\partial t}(s+\delta, s) - \frac{\partial G}{\partial t}(s-\delta, s) &= \lim_{\delta \to 0^+} c^{-1}[y_1(s)y_2^{'}(s+\delta) - y_1^{'}(s-\delta)y_2(s)] \\ &= c^{-1} W(y_1, y_2)(s) \\ &= \frac{1}{a_0(s)} \end{aligned} \tag{4.20}$$

(3)G is in the kernel space of L, namely

$$\begin{aligned} LG(\cdot, s) &= c^{-1} y_1(s) L y_2 \quad (\text{or } c^{-1}(Ly_1)y_2(s)) \\ &= 0 \end{aligned} \tag{4.21}$$

Proposition 4.4. *The operator* $T : C(I) \to C^2(I)$

$$Tf(t) := \int_I G(t,s)f(s)ds \in C^2(I) \tag{4.22}$$

solves the equation (4.1), *i.e.* $LTf = f$.

Proof.

$$\begin{aligned}(Tf)^{'}(t) &= \int_a^t G_t(t,s)f(s)ds + G(t,t^-)f(t^-) \\ &+ \int_t^b G_t(t,s)f(s)ds - G(t,t^+)f(t^+) \\ &= \int_a^t G_t(t,s)f(s)ds + \int_t^b G_t(t,s)f(s)ds\end{aligned} \tag{4.23}$$

due to the continuity of G.

$$\begin{aligned}(Tf)^{''}(t) &= \int_a^t G_{tt}(t,s)f(s)ds + G_t(t,t^-)f(t^-) \\ &+ \int_t^b G_{tt}(t,s)f(s)ds - G_t(t,t^+)f(t^+) \\ &= \int_a^t G_{tt}(t,s)f(s)ds + \int_t^b G_{tt}(t,s)f(s)ds + \frac{f(t)}{a_0(t)}\end{aligned} \tag{4.24}$$

by the same argument as (4.20).

Hence

$$\begin{aligned}L(Tf)(t) &= a_0(t)(Tf)^{''}(t) + a_0^{'}(t)(Tf)^{'}(t) + a_2(t)Tf(t) \\ &= (\int_a^t + \int_t^b)L_tG(t,s)f(s)ds + f(t) \\ &= f(t)\end{aligned} \tag{4.25}$$

due to (4.21).

If $g \in C^2(I)$ satisfies (4.13)

$$\begin{aligned}
TLg(t) &= (\int_a^t + \int_t^b)[\frac{d}{ds}(a_0(s)\frac{dg(s)}{ds}) + a_2(s)g(s))]G(t,s)ds \\
&= a_0(s)\frac{dg(s)}{ds}G(t,s)|_a^b - a_0(s)g(s)G_s(t,s)|_a^t - a_0(s)g(s)G_s(t,s)|_t^b \\
&+ (\int_a^t + \int_t^b)L_sG(t,s)g(s)ds \\
&= a_0(s)g(s)G_s(t,s)|_{t^-}^{t^+} + a_0(s)[g^{'}(s)G(t,s) - g(s)G_s(t,s)]|_a^b \\
&= g(t)
\end{aligned} \tag{4.26}$$

where we have used (4.20) and the fact that both g and G satisfy (4.13).

□

Here I give a lemma and leave its proof as an exercise.

Lemma 4.5. *If A is a self-adjoint compact operator on a Hilbert space $\mathscr{H}$, then there exists an eigenvalue λ such that $|\lambda| = \|A\|$.*

Remark 4.6. T defined in (4.22) is equicontinuous and uniformly bounded due to the fact that $G(\cdot, s)$ is uniformly continuous on I, or equivalently T is compact as we have shown in Lemma 4.8. Thus Arzelà Ascoli Theorem applies and the Lemma 4.5 holds for T. To prove $\|A\| = \sup\limits_{\|x\|=1} |\langle Ax, x\rangle|$, no matter we deal with a Hilbert space $\mathscr{H}$ or $C(I)$, we only have to show that $2Re\langle Tu, v\rangle \leqslant (\|u\|^2 + \|v\|^2) \sup |\langle Tu, u\rangle|$ and then replace v with $Tu/\|Tu\|$.

The following theorem describes the distribution of eigenvalues of a linear, self-adjoint compact operator.

Theorem 4.7. *(Hilbert-Schmidt) Let A be a linear, self-adjoint, compact operator on a Hilbert space $\mathscr{H}$ with $\dim(\mathscr{H}) = \infty$. All eigenvalues of A are real and can be ordered so that*

$$|\lambda_{n+1}| \leqslant |\lambda_n|, \quad \lim_{n\to\infty} \lambda_n \to 0 \tag{4.27}$$

Furthermore, eigenvectors $\{\varphi_n\}_{n\in\mathbb{N}}$ can be chosen to be an ONB of $Ran(A)$. In particular, when $\ker(A) = 0$, each element $h \in \mathscr{H}$ has an expansion $h = \sum\limits_n \lambda_n \langle h, \varphi_n\rangle \varphi_n$.

Proof. By Lemma 4.5, there exists φ_1 such that $\|\varphi_1\| = 1$ and $A\varphi_1 = \lambda_1\varphi_1$ with $|\lambda_1| = \|A\|$. Let $H_1 = \overline{span\{\varphi_1\}}^{\perp}$, then $A|_{H_1}$ is another self-adjoint compact operator with operator norm $\|A|_{H_1}\| \leqslant \|A\|$. It has an eigenvector φ_2 such that $A|_{H_1}\varphi_2 = \lambda_2\varphi_2$ with $\lambda_2 = \|A|_{H_1}\| \leqslant \lambda_1$. Continuing the same argument gives us a sequence of eigenvalues such that $|\lambda_1| \geqslant |\lambda_2| \geqslant \cdots$
Assume eigenvalues of A are bounded away from zero, namely $|\lambda_n| \geqslant b > 0$. Then the fact that $\{b\varphi_n\}_{n\in\mathbb{N}} \subset A(B_H(0,1))$ has no convergent subsequence(φ_n orthonormal) contradicts the compactness of A.
Finally, if $x \in \overline{span\{\varphi_1, \varphi_2, \cdots\}}^{\perp}$, then $\|Ax\| \leqslant \lambda_n\|x\|$ for each $n \in \mathbb{N}$, hence $Ax = 0$. If $\ker(A) = \{0\}$, then $H = \overline{span\{\varphi_1, \varphi_2, \cdots\}}$. □

Lemma 4.8. *If $k \in L^2_w(I \times I)$, then integral operator $K_w : L^2_w(I) \to L^2_w(I)$, $K_w f(t) := \int_I k(t,s)f(s)w(s)ds$ is compact.*

Proof. Let $\{\phi_i\}$ be an ONB of $L^2_w(I)$ and suppose $k(x,y) = \sum\limits_{i,j=1}^{\infty} k_{ij}\phi_i(x)\phi_j(y)$. Then $\|k\|^2_{L^2_w(I\times I)} = \sum\limits_{ij} |k_{ij}|^2$. For any $f \in L^2_w(I)$, denote by $f_j = \langle f, \phi_j\rangle_w$, we have

$$\begin{aligned}
\langle K_w f, K_w f\rangle_w &= \int_I \left(\sum_{i,j=1}^{\infty} k_{ij}\phi_i(t)f_j\right)\left(\sum_{i',j'=1}^{\infty} k_{i'j'}\phi_{i'}(t)f_{j'}\right) w(t)dt \\
&= \sum_i \sum_j \sum_{j'} k_{ij}k_{ij'}f_j f_{j'} \\
&\leqslant \sqrt{\sum_{i,j,j'} (k_{ij}f_{j'})^2 \sum_{i,j,j'} (k_{ij'}f_j)^2} \\
&= \|f\|^2_w \|k\|^2_w
\end{aligned} \tag{4.28}$$

namely Kf belongs to L^2_w.

Given any $n \in \mathbb{N}$, $K_n f := \sum\limits_{i,j=1}^{n} k_{ij}\langle f, \phi_i\rangle\phi_j$ is a bounded operator with finite dimensional range, hence compact. $\|K_w - K_n\|^2 = \sum\limits_{i,j=n+1}^{\infty} |k_{ij}|^2 \to 0$, namely $K_n \to K_w$ in operator norm, thus we claim that K is compact. Indeed, let $\{f_m\}$ be a bounded sequence, Arzelá-Ascoli argument tells us that there exists a subsequence $\{f_{m_j}\}$ such that, for each $n \in \mathbb{N}$, $K_n f_{m_j}$ is convergent, hence $\{K_w f_{m_j}\}$ a Cauchy sequence. Thus the completeness of $L^2_w(I)$ gives us a convergent subsequence $\{K_w f_{m_j}\}$.

□

Let

$$T_w f = \int_I G(t,s) f(s) ds \tag{4.29}$$

By Proposition 4.4, we see that

$$\frac{1}{w} L T_w f = f \tag{4.30}$$

for positive w. Note that if $y_i \in L^2_w(I)(i = 1, 2)$, then $G \in L^2_w(I \times I)$, hence T_w is linear, self-adjoint and compact. Since the eigenvalues of T_w goes to zero, as a left-inverse operator, the Sturm-Liouville operator has eigenvalues $\{\lambda_n\}$ which is increase and goes to infinity. When y_n is some eigenfunction of T_w corresponding to eigenvalue $\tilde{\lambda}_n > 0$, then $Ly_n = \frac{1}{\tilde{\lambda}_n} w y_n$. Thus we have proved the following result that I intend to give in this section, which allows us to study the functions on I in terms of various kinds of orthogonal polynomials with different properties.

Theorem 4.9. *The eigenfunctions of SL-problem* (4.12) *form a complete ONB for $L^2_w(I)$, with w strictly positive.*

Before I proceed to exhibit examples of orthogonal polynomials, let us have a brief look at the zeros of the solution functions, which is a topic that attracted several generations of mathematicians. The zeros of a non-trivial solution u of (4.12) in (a, b) are isolated from the simple observation that, $u(x_0)$ and $u^{'}(x_0)$ cannot be zero at the same time, hence u must increase or drop in a neighborhood of the zero x_0.

Assume $I^{'} \subset I$ is an subinterval on which a solution u_1 of (4.12) does not vanish. If u_2 is another solution that is independent of u_1, then $u_1(a_0 u_2^{'}) - u_2(a_0 u_1^{'}) = a_0 W(u_1, u_2)$ is a non-zero constant $\hat{c}$, hence we have

$$(\frac{u_2}{u_1})^{'}(x) = \frac{\hat{c}}{a_0(x) u_1^2(x)}, \text{ for } x \in I^{'} \text{a.e.} \tag{4.31}$$

Integrating on $(a^{'}, x) \subset I^{'}$ yields

$$\frac{u_2}{u_1}(x) = \frac{u_2}{u_1}(a^{'}) + \int_{a'}^{x} \frac{\hat{c}}{a_0 u_1^2} \tag{4.32}$$

Suppose u_1 and u_2 are two linearly independent solutions. When x_1 and x_2 are successive zeros of a solution u_1, from $W(u_1, u_2) \neq 0$, we shall have $u_1^{'}(x_1) u_2(x_1)$ and $u_1^{'}(x_2) u_2(x_2)$ are either both positive or negative. Since $u_1^{'}(x_1)$ and $u_1^{'}(x_2)$ have opposite sign, $u_2(x_1) u_2(x_2) < 0$. There must exist exactly one point $x_3 \in (x_1, x_2)$ such that $u_2(x_3) = 0$, for otherwise it would contradicts the assumption that x_1 and x_2 are successive zeros of u_1. We conclude that

Proposition 4.10. *The zeroes of two linearly independent solutions of* (4.12) *must intertwine with each other.*

Now we arrive at some famous polynomials coming as the eigenfunctions of the Sturm-Liouville eigenvalue problems.

Example 4.11. (Laguerre Polynomials)
Laguerre polynomials

$$L_{n,\alpha}(x) = \frac{e^x x^{-\alpha}}{n} \frac{d^n}{dx^n} x^{n+\alpha} e^{-x} \tag{4.33}$$

are non-singular solutions of

$$xu^{''} + (1+\alpha-x)u^{'} + nu = 0,\ 0 < x < \infty \tag{4.34}$$

with $\alpha > -1$ and n non-negative integer; especially when $\alpha = 0$ it is equivalent to

$$(xe^{-x}u^{'})^{'} + ne^{-x}u = 0 \tag{4.35}$$

Obviously in this example $a_0(x)$ vanishes at $x = 0$. Those polynomials are also eigenfunctions of the Sturm-Liouville operator in (4.35) with respect to eigenvalues $\lambda_n = n(n+1)$, with the orthogonality property that $\langle L_{n,\alpha}, L_{m,\alpha}\rangle_{e^{-x}x^{\alpha}} = \delta_{n,m}$ in $L^2_{e^{-x}x^{\alpha}}(0,\infty)$. Its series expression is

$$L_{n,\alpha}(x) = \sum_{k=0}^{n} \frac{\Gamma(n+\alpha+1)}{\Gamma(k+\alpha+1)} \frac{(-x)^k}{k!(n-k)!} \tag{4.36}$$

Besides, they satisfy the recurrence relation

$$nL_{n,\alpha}(x) = (2n+\alpha-1-x)L_{n-1,\alpha}(x) - (n+\alpha-1)L_{n-2,\alpha}(x) \tag{4.37}$$

Example 4.12. (Legendre Polynomials)
When we solve the Laplace equation in $\mathbb{R}^3$

$$\frac{\partial}{\partial r}(r^2\frac{\partial u}{\partial r}) + \frac{1}{\sin^2\varphi}\frac{\partial^2 u}{\partial\theta^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\varphi}(\sin\varphi\frac{\partial u}{\partial\varphi}) = 0 \tag{4.38}$$

by assuming that the solution u is independent of θ and using the method of separation of variables, we obtain the two ODEs with respect to φ and r respectively, namely

$$r^2v^{''} + 2rv^{'} - \lambda v = 0 \tag{4.39}$$

and

$$\frac{1}{\sin\varphi}(\sin\varphi w^{'})^{'} + \lambda w = 0 \tag{4.40}$$

A change of variable $u = \cos\varphi$ gives the Legendre's equation

$$(1-x^2)u^{''} - 2xu^{'} + \lambda_n u = 0, \ -1 < x < 1 \tag{4.41}$$

which has singular points $x = 1$, where $\lambda_n = n(n+1)$, $n \in \mathbb{N}$ and the corresponding eigenfunctions P_n are Legendre polynomials, having the orthogonality $\langle P_n, P_m \rangle = \frac{2}{2n+1}\delta_{m,n}$ in $L^2(-1,1)$ if we assume $P_n(1) = 1$.

From (1.19) it can be derived that

$$(1-x^2)\frac{d}{dx}P_l(x) = l(P_{l-1}(x) - xP_l(x)) \tag{4.42}$$

and that

$$(1-x^2)\frac{d}{dx}P_l(x) = (l+1)(xP_l(x) - P_{l+1}(x)) \tag{4.43}$$

Taking derivative on both sides of the above equations and using (4.41) give

$$lP_l(x) = x\frac{d}{dx}P_l(x) - \frac{d}{dx}P_{l-1}(x) \tag{4.44}$$

and

$$(l+1)P_l(x) = -x\frac{d}{dx}P_l(x) + \frac{d}{dx}P_{l+1}(x) \tag{4.45}$$

Example 4.13. (Hermite Polynomials)
Probably the most well known polynomials in Fourier analysis is the Hermite polynomials, due to the fact that they constitute eigenfunctions of the Fourier transform on $\mathbb{R}$, namely $\hat{H}_n = (-i)^n H_n$. They solve the Hermite's equation:

$$u^{''} - 2xu^{'} + 2nu = 0, \ -1 < x < 1 \tag{4.46}$$

or

$$(e^{-x^2}u^{'})^{'} + 2ne^{-x^2}u^{'} = 0, \ -1 < x < 1 \tag{4.47}$$

Therefore, they are at the same time eigenfunctions $H_n = (-1)^n e^{x^2}\frac{d^n}{dx^n}e^{-x^2}$ of the SL operator in (4.47) with respect to eigenvalues $\lambda_n = 2n$. By induction, $H_n(x) = (2x)^n + (-1)^n p(x)$ where the polynomial degree of $p(x)$ is less than n. In fact, it has the series expression

$$H_n(x) = \sum_{k=0}^{\llcorner n/2 \lrcorner} \frac{(-1)^k n!}{k!(n-2k)!}(2x)^{n-2k} \tag{4.48}$$

Setting $H_{-1} \equiv 0$, the recurrence relation reads

$$H_n(x) = 2xH_{n-1}(x) - 2nH_{n-2}(x) \tag{4.49}$$

Hermite polynomials have the orthogonality property in $L^2_{e^{-x^2}}(\mathbb{R})$

$$\langle H_n, H_m \rangle_{e^{-x^2}} = \int_{-\infty}^{\infty} \frac{d^n}{dx^n} H_m(x) e^{-x^2} dx = 2^n n! \sqrt{\pi} \delta_{m,n} \tag{4.50}$$

where we used the fact that $\frac{d^n}{dx^n} H_m(x) = 0$ for polynomial degree $m < n$. Besides, it is easy to verify that Hermite polynomials' generating function is $e^{2xt-t^2} = \sum\limits_{n=0}^{\infty} \frac{1}{n!} H_n(x) t^n$.

Example 4.14. (Bessel Functions)
Bessel's equation takes the form

$$x^2 u^{''} + xu^{'} + (x^2 - \nu^2)u = 0 \tag{4.51}$$

where ν is a nonnegative parameter. It obviously has singular point at $x = 0$. Here the SL eigenvalue problem generates non-polynomial eigenfunctions J_ν, called the Bessel functions.

In fact, by assuming $u = x^s \sum\limits_{k=0}^{\infty} c_k x^k$ and collecting the coefficients of the powers $x^s, x^{s+1}, \cdots$, we get $s = \nu$ and $c_k = -\frac{1}{k(k+2\nu)} c_{k-2}$, hence $c_{2m} = \frac{(-1)^m}{2^{\nu+2m} m! \Gamma(\nu+m+1)} x^{2m}$ and $c_{2m+1} = 0$. Let

$$J_\nu(x) := (\frac{x}{2})^\nu \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(\nu + m + 1)} (\frac{x}{2})^{2m} \tag{4.52}$$

then

$$\lim_{x \to 0^+} J_\nu(x) = \begin{cases} 1 & \nu = 0 \\ 0 & \nu > 0 \end{cases} \tag{4.53}$$

tells us that J_ν is well defined for $\nu \geqslant 0$. For $t = -\nu$, we can similarly define

$$J_{-\nu}(x) := (\frac{x}{2})^{-\nu} \sum_{m=0}^{\infty} \frac{(-1)^m}{m! \Gamma(\nu + m + 1)} (\frac{x}{2})^{2m} \tag{4.54}$$

but it is not necessarily bounded at $x = 0$. In fact, when $\nu = n$, $J_{-\nu} = (-1)^n J_\nu$; but for $\nu \in \mathbb{R}^+ \setminus \mathbb{N}_0$, $\lim\limits_{x \to 0^+} |J_{-\nu}| = \infty$. Thus we have arrived at the following conclusion.

Lemma 4.15. *Bessel functions J_ν and $J_{-\nu}$ are linearly independent iff $\nu \in \mathbb{R}^+ \setminus \mathbb{N}_0$.*

If we take $a = 0$ and $b < \infty$, the scaled Bessel Functions form an orthogonal and complete basis in $L^2_x(0, b)$, namely

$$\langle J_n(\sqrt{\lambda_j}x), J_n(\sqrt{\lambda_k}x)\rangle_x = \delta_{j,k} \tag{4.55}$$

where the eigenvalue $\lambda_k = (\frac{x_{nk}}{b})^2$ and x_{nk} is the k-th zero of J_n, i.e. $0 < x_{n1} < x_{n2} < \cdots < x_{nk} < \cdots$, coming as the eigenvalue of the scaled Bessel's equation

$$xu^{''} + u^{'} + (\lambda x - \frac{\nu^2}{x})u = 0 \tag{4.56}$$

We are not going to prove those, but rather refer to [13] for a thorough and exquisite exposition of the properties of Bessel functions.

4.2 Spectrum of discrete Laplacian

Given a finite subgraph H of G, analogues of the length of boundary of a submanifold H in the continuous setting could be the number of joint edges of vertex set in H with its complement vertex set

$$|E(\partial_v H)| = |\{xy \in E(G) : x \in V(H), y \in V(G)\backslash V(H)\}| \tag{4.57}$$

or probably more naturally defined as the number of edges between faces in H and its complement part, namely

$$|E(\partial_f H)| = |\{xy \in E(H) \cap E(F(G)\backslash F(H))\}| \tag{4.58}$$

Similarly, the surface area of H can be measured either by $A(H) = \sum_{v \in V(H)} \deg(v)$ or by the number of faces $|F(H)|$ in H. These lead to two different isoperimetric constants measuring the ratio of length and area, that are closely connected to the curvature, namely

$$\alpha(G) = \inf_{0<|H|\leqslant\frac{1}{2}|V(G)|} \{|E(\partial_v H)|/A(H)\} \tag{4.59a}$$

$$\alpha^*(G) = \inf_{0<|H|\leqslant\frac{1}{2}|V(G)|} \{|E(\partial_f H)|/|F(H)|\} \tag{4.59b}$$

where $\alpha(G)$ is called Cheeger constant.

Let $C^0(G)$ be the space of real valued functions of vertices with inner product

$$(g_1, g_2) = \sum_{v \in V(G)} g_1(v) g_2(v)$$

and $\{\mathbb{R}\}$ be the set of real constant functions on G. The set of 1-forms $C^1(G)$ is defined as functions on oriented edges satisfying $\omega([x,y]) = -\omega([y,x])$ endowed with inner product

$$(\omega_1, \omega_2) = \sum_e \omega_1(e)\omega_2(e)$$

In particular, there is $dg([x,y]) = \nabla_{xy} g = g(y) - g(x)$ for any $g \in C^0(G)$. Similarly, we can define $C^0(G,H) \subset C^0(G)$ consisting of those functions vanishing on subcomplex H. It follows that

$$\begin{aligned}(dg_1, dg_2) &= \sum_{[x,y]} (g_1(y) - g_1(x)) (g_2(y) - g_2(x)) \\ &= \sum_x \sum_{y \in \mathcal{N}(x)} (g_1(x) - g_1(y)) g_2(x)\end{aligned}$$

where $[x,y]$ run over all the possible edges in $E(G)$ with an arbitrarily chosen direction on each of them. Thus the Laplacian on G can be defined as

$$\Delta_G g(x) = -d^* dg(x) = \sum_{y \in \mathcal{N}(x)} (g(y) - g(x)) = \sum_{y \in \mathcal{N}(x)} g(y) - \deg(x) \cdot g(x) \quad (4.60)$$

The restricted Laplacian Δ_H on subcomplex H can be defined correspondingly as Δ_G restricted on subspace $C^0(G,H)$.

If G is an infinite connected graph, it was proved in [75] under the geometric assumptions that

$$\begin{aligned}&a.\ \deg(v) \text{ is uniformly bounded above by some constant } p \\ &b.\ \text{there exists positive constant } \gamma \text{ such that } \gamma V(H) \leqslant V(\partial H)\end{aligned} \quad (4.61)$$

for any finite subcomplex $H < G$, the minimal positive eigenvalue of $-\Delta_H$(and $-\Delta_G$) has lower bound $\frac{\gamma^2}{2p}$(respectively).

Theorem 4.16. *If $H \setminus \partial H$ is connected, then the minimal eigenvalue of $-\Delta_H$ in $\sigma(-\Delta_H) \backslash \{0\}$ satisfies*

$$\lambda_{min} = \inf_{g \in C^0(H, \partial H) \backslash \{\mathbb{R}\}} \frac{(-\Delta_H g, g)}{(g,g)} \geqslant \frac{\gamma^2}{2p} \quad (4.62)$$

and the infimum can be achieved by a positive, subharmonic and nonconstant function h. As H asymptotically approach G, if follows that the minimal positive eigenvalue of $-\Delta_G$ satisfies (4.62) *with a positive, subharmonic and nonconstant eigenfunction.*

Proof. The idea of the proof is based on estimation of the quantity

$$\sum_e \left|d(g^2)(e)\right| = \sum_{[x,y]} \left|g^2(x) - g^2(y)\right|$$

where $g \in C^0(H, \partial H)$ while $[x, y]$ run over the set of edges. On the one hand, there is

$$\begin{aligned}\sum_e \left|d(g^2)(e)\right| &\leqslant \left(\sum_{[x,y]} |g(x) + g(y)|^2\right)^{1/2} \cdot \left(\sum_{[x,y]} |g(x) - g(y)|^2\right)^{1/2} \\ &\leqslant \sqrt{2}\left(\sum_{[x,y]} (g^2(x) + g^2(y))\right)^{1/2} \cdot (dg, dg)^{1/2} \\ &\leqslant \sqrt{2p}(g, g)^{1/2}(dg, dg)^{1/2}\end{aligned} \tag{4.63}$$

On the other hand, note that for any $g \in C^0(H, \partial H) \setminus \{\mathbb{R}\}$

$$\frac{(g(x) - g(y))^2}{(g, g)} \geqslant \frac{(|g(x)| - |g(y)|)^2}{(|g|, |g|)} \geqslant \inf_{g \notin \{\mathbb{R}\}} \frac{(-\Delta_H g, g)}{(g, g)}$$

hence if g is an eigenfunction corresponding to the smallest positive eigenvalue λ_{min}, so is $|g|$. In fact, by the connectness assumption, $g = |g|$, for otherwise g has negative value, then there exists x and its neighboring point y such that $(g(x) - g(y))^2 > (|g(x)| - |g(y)|)^2$, which leads to the absurdity that $\lambda_{min} > \lambda_{min}$. There cannot be a point $x \in H \setminus \partial H$ such that $g(x) = 0$, for otherwise $-\lambda_{min} g(x) = \Delta_H g(x) = \sum_{y \in \mathcal{N}(x)} (g(y) - g(x)) \geqslant 0$, hence g is identically zero on H, a contradiction as well. Thus we see the positivity of the eigenfunction g.

Now suppose the values set of g on H is ordered as $0 = \beta_0 < \beta_1 < \cdots < \beta_N$.

We have

$$\begin{aligned}
&\sum_{g(x)=\beta_i} \sum_{k=1}^{i} \sum_{\{y\in\mathcal{N}(x):g(y)=\beta_{i-k}\}} \left(g^2(x)-g^2(y)\right) \\
&= \sum_{g(x)=\beta_i} \sum_{k=1}^{i} \sum_{\{y\in\mathcal{N}(x):g(y)=\beta_{i-k}\}} \sum_{t=i-k+1}^{i} \left(\beta_t^2-\beta_{t-1}^2\right) \\
&= \sum_{g(x)=\beta_i} \sum_{t=1}^{i} \sum_{k\geqslant i-t+1} \sum_{\{y\in\mathcal{N}(x):g(y)=\beta_{i-k}\}} \left(\beta_t^2-\beta_{t-1}^2\right) \\
&= \sum_{g(x)=\beta_i} \sum_{t=1}^{i} \left(\beta_t^2-\beta_{t-1}^2\right) \sum_{k\geqslant i-t+1} |\{y\in\mathcal{N}(x):g(y)=\beta_{i-k}\}|
\end{aligned}$$

Thus

$$\begin{aligned}
\sum_e \left|d(g^2)(e)\right| &= \sum_{i=1}^{N} \sum_{g(x)=\beta_i} \sum_{t=1}^{i} \left(\beta_t^2-\beta_{t-1}^2\right) \sum_{k\geqslant i-t+1} |\{y\in\mathcal{N}(x):g(y)=\beta_{i-k}\}| \\
&= \sum_{t=1}^{N} \left(\beta_t^2-\beta_{t-1}^2\right) \sum_{g(x)\geqslant\beta_t} |\{y\in\mathcal{N}(x):g(y)\leqslant\beta_{t-1}\}|
\end{aligned} \tag{4.64}$$

$\Delta_H g(x) < 0$ indicates that there is at least a neighboring point y such that $g(y) < g(x)$. By (4.61) there is

$$\begin{aligned}
\sum_e \left|d(g^2)(e)\right| &\geqslant \sum_{t=1}^{N} \left(\beta_t^2-\beta_{t-1}^2\right) |\partial\{x:g(x)\geqslant\beta_t\}| \\
&\geqslant \gamma \sum_{t=1}^{N} \left(\beta_t^2-\beta_{t-1}^2\right) |\{x:g(x)\geqslant\beta_t\}| \\
&= \gamma \left[\beta_N^2|\{g(x)=\beta_N\}| + \sum_{t=1}^{N-1} \beta_t^2(|\{g(x)\geqslant\beta_t\}| - |\{g(x)\geqslant\beta_{t+1}\}|)\right] \\
&\geqslant \gamma(g,g)
\end{aligned} \tag{4.65}$$

Combining (4.63) and (4.65) gives the desired result (4.62). Besides, if λ_{min} is obtained by function g, then $-\Delta_H g = \lambda_{min} g > 0$, namely g is subharmonic and nonconstant. g being positive and subharmonic also gives that Harnack inequality

$$\frac{1}{\deg(y)}g(x) \leqslant g(y) \leqslant \deg(y)g(x) \tag{4.66}$$

which follows immediately from definition.

Choose an arbitrary point in x_0 in G, and we can form a finite subgraph H_n including vertices connected to x_0 by at most n edges from $E(G)$. The second part of the assertion comes as the limit of the smallest eigenvalue λ_n and associated eigenfunctions of Δ_{H_n}. In fact, let $g_n \in C^0(H_n, \partial H_n) \setminus \{\mathbb{R}\}$ such that $g_n(x_0) = 1$, where ∂H_n consists of points in H_n that have at least one neighboring point not in H_n. Notice that $C^0(H_n, \partial H_n) \subset C^0(H_{n+1}, \partial H_{n+1})$, hence $\lambda_n = \min\limits_{g \in C^0(H_n, \partial H_n)\setminus\{\mathbb{R}\}} \frac{(-\Delta_H g, g)}{(g,g)} \geqslant \lambda_{n+1} \geqslant \frac{\gamma^2}{2p}$ and there exists $\lambda \geqslant \frac{\gamma^2}{2p}$ as the limit of λ_n.

Denote by g_n the eigenfunction corresponding to λ_n. Then by (4.66), if vertex $x \in G$ is of s-edge distance to x_0, there is

$$\frac{1}{\deg(y)^{\min\{n,s\}}}g_n(x) \leqslant g_n(y) \leqslant \deg(y)^{\min\{n,s\}}g_n(x) \tag{4.67}$$

and if $n < s$, $g_n(x) = 0$. Therefore $\{g_n(x)\}_{n \geqslant 0}$ is a bounded sequence. A diagonal argument gives a subsequence such that $\lim\limits_{k\to\infty} g_{n_k}(x)$ exists for all vertices in G, and we define through pointwise value a function h. Clearly $h(x_0) = \lim\limits_{n\to\infty} g_n(x_0) = 1$, and $-\Delta_G h = \lambda h \geqslant 0$. In fact, for any x in G, (4.67) gives that $h(x)$ is strictly positive, namely h is positive and nonconstant.

□

Remark 4.17. In the original paper [75], it was somehow miscalculated that $\sum\limits_e \left|d(g^2)(e)\right| = \sum\limits_{t=1}^{N} \left(\beta_t^2 - \beta_{t-1}^2\right) \left|\partial\{x : g(x) = \beta_t\}\right|$, although the final conclusion remains correct. In fact, we see in (4.65) that this holds only with an inequality, and the precise expression for $\sum\limits_e \left|d(g^2)(e)\right|$ is given in (4.64), which is exactly $\sum\limits_{t=1}^{N} (\beta_t^2 - \beta_{t-1}^2) \left|E(\partial_v L_t)\right|$ with $L_t := \{x : g(x) \geqslant \beta_t\}$. Probably noticed this imprecision, in a later work[76] the same author changed the setting for Laplacian and used a different assumption (4.59) in replacement of (4.61). Thus the same argument as in (4.65) applies and gives the estimation

$$\begin{aligned}\sum_e \left|d(g^2)(e)\right| &\geqslant \alpha \sum_{t=1}^{N} A(L_t) \\ &= \alpha(\beta_t^2 - \beta_{t-1}^2) \sum_{x \in L_t} \deg(x) \\ &= \alpha \sum_{t=1}^{N-1} \beta_t^2 \sum_{g(x)=\beta_t} \deg(x) + \alpha\beta_N^2 \deg(x) \\ &= \alpha \sum_{x \in H} \deg(x) g^2(x) \\ &= \alpha(g,g)_w\end{aligned} \tag{4.68}$$

Meanwhile (4.63) gives $\sum_e \left|d(g^2)(e)\right| \leqslant \sqrt{2}(g,g)_w^{1/2}(dg,dg)^{1/2}$. Together it gives the minimal positive eigenvalue estimation

$$\lambda^*_{\min} \geqslant \frac{\alpha^2}{2}$$

for weighted Laplacian Δ_w. The positivity of α was given in same work under the assumption that $\deg(x) \geqslant 7$ for all vertices $x \in V$, for that there is $\alpha(G) \geqslant \frac{1}{78}$. The other situations are verified as the main result of [61]. In fact, it is proved there that $\kappa(v) < 0$ for every $v \in V(G)$ implies $\alpha^*(G) > 0$; while $\chi(f) < 0$ for every face in $F(G)$ implies $\alpha(G) > 0$. Here

$$\chi(f) = 1 - \frac{|f|}{2} + \sum_{v \in f} \frac{1}{\deg(v)} \tag{4.69}$$

is the *Euler-characteristic* with $|f|$ the number of vertices contained in f, and

$$\kappa(v) = 1 - \frac{\deg(v)}{2} + \sum_{\{f: v \in f\}} \frac{1}{|f|} \tag{4.70}$$

is the *combinatorical curvature* at a vertex v. The discrete Gauss-Bonnet theorem says

$$\sum_v \kappa(v) = 2 - \text{genus} \tag{4.71}$$

where the genus as usual is defined as the maximal numbers of nonintersecting simple closed curves that can be drawn on a surface without separating it. For a proof of this fundamental result I refer the readers to [50].

Theorem 4.18. *On a weighted finite graph without loops, there is*

$$\frac{\alpha^2(G)}{2} \leqslant \lambda^*_{\min} \leqslant \min\{2\alpha(G), \frac{V(G)-2}{|V(G)|-1}\}$$

and

$$\max\{2-2\alpha(G), \frac{|V(G)|}{|V(G)|-1}\} \leqslant \lambda^*_{\max} \leqslant 2 - \frac{\alpha^2(G)}{2}$$

Proof. The lower bound for $\lambda^*_{\min}$ has been given in the previous remark. In the upper bound aspect, given any nonempty $W \subset V(G)$, by letting

$$g^{\sharp}_W(x) = \begin{cases} 1 & x \in W \\ -\frac{A(W)}{A(W^c)} & x \in W^c \end{cases} \tag{4.72}$$

we get

$$\begin{aligned} \frac{(-\Delta_w g^{\sharp}_W, g^{\sharp}_W)_w}{(g^{\sharp}_W, g^{\sharp}_W)_w} &= \frac{\sum\limits_{x\in\partial W} \left(1 + A(W)/A(W^c)\right)^2 A(x)}{A(W) + \sum\limits_{x\in W^c} A(x)A^2(W)/A^2(W^c)} \\ &= \frac{A(\partial W)A(V)}{A(W^c)A(W)} \\ &\leqslant 2\alpha \end{aligned} \tag{4.73}$$

Let δ_v be orthonormal basis on $V(G)$ such that $\delta_v(v^{'}) = 1$ for $v^{'} = v$, and zero otherwise. The trace of the discrete Laplacian is

$$\begin{aligned} \sum_{i=0}^{N} \lambda^*_i &= \sum_{v\in V} \frac{(\Delta_w \delta_v, \delta_v)}{(\delta_v, \delta_v)} \\ &= \frac{1}{2}\sum_{v\in V} \frac{\sum\limits_{u^{'},v^{'}} \left(\delta_v(u^{'}) - \delta_v(v^{'})\right)^2 \delta_{v^{'},u^{'}}}{\sum\limits_{v^{'}} \delta_v^2(v^{'})A(v^{'})} \\ &= |V| \end{aligned} \tag{4.74}$$

Therefore

$$\lambda^*_N = \lambda^*_{\max} \geqslant \frac{|V|}{N} \tag{4.75}$$

where $N+1 = |V|$. Meanwhile, let f_N be the eigenfunction of the largest eigenvalue λ_N^*, we have

$$\begin{aligned}
\lambda_1^* + \lambda_N^* &\leqslant \frac{(\Delta_w |g_N|, |g_N|)_w}{(|g_N|, |g_N|)_w} + \frac{(\Delta_w g_N, g_N)_w}{(g_N, g_N)_w} \\
&= \frac{\sum\limits_{v,v' \in V(G)} \left[\left(|g_N(v)| - |g_N(v')|\right)^2 + \left(g_N(v) - g_N(v')\right)^2\right] A_{vv'}}{2 \sum\limits_{v \in V(G)} g_N^2(v) A(v)} \\
&\leqslant \frac{\sum\limits_{v,v' \in V(G)} \left(g_N^2(v) + g_N^2(v')\right) A_{vv'}}{\sum\limits_{v \in V(G)} g_N^2(v) A(v)} \\
&= 2
\end{aligned} \tag{4.76}$$

hence $\lambda_1^* \leqslant 2 - \frac{|V|}{N} = \frac{V-2}{V-1}$ while $\lambda_N^* \leqslant 2 - \frac{\alpha^2}{2}$.

□

4.3 Other representations on the sphere

Rotation group is closely related to quaternions, which provides us with an alternative and convenient way to represent points and compute on the sphere. Let $\mathbb{H}$ denote the set of quaternions consisting of elements like

$$\mathbf{q} = q_0 + q_1\mathbf{i} + q_2\mathbf{j} + q_3\mathbf{k} = q_0 + \mathfrak{Vec}(\mathbf{q}) \tag{4.77}$$

which can be represented as $\mathbf{q} = \|\mathbf{q}\|(\cos\theta + \hat{n}\sin\theta)$, analogous to complex numbers, where $\|\mathbf{q}\| = \sqrt{\mathbf{q}\bar{\mathbf{q}}}$ and $\hat{n} = (n_1, n_2, n_3)$ is the unit vector in the direction of the vector part of the quaternion $\mathbf{q}$, i.e. $\hat{n}\|\mathbf{q} - q_0\| = \mathbf{q} - q_0$. Therefore a unit quaternion means a rotation about the axis $\hat{n}$ by the angle 2θ. $\mathbb{H}$ is a division algebra, and the inverse of a nonzero element $\mathbf{p}$ is $\mathbf{p}^{-1}/\|\mathbf{p}\|^2$. We denote the set of all unit quaternion by $\mathbb{B}_{\mathbb{H},1}$. Whenever two quaternions $\mathbf{p}$,$\mathbf{q}$ are conjugate with each other in the sense that there exists $\mathbf{r} \in \mathbb{H}$ such that $\mathbf{p} = \mathbf{r}\mathbf{q}\mathbf{r}^{-1}$, they have exactly the same norm and the same real part. In fact, we have the relation

$$\begin{aligned}\mathfrak{Re}(\mathbf{pq}) &= p_0q_0 - \mathfrak{Vec}(\mathbf{p}) \cdot \mathfrak{Vec}(\mathbf{q})\\ \mathfrak{Vec}(\mathbf{pq}) &= p_0\mathfrak{Vec}(\mathbf{q}) + q_0\mathfrak{Vec}(\mathbf{p}) + \mathfrak{Vec}(\mathbf{p}) \times \mathfrak{Vec}\mathbf{q}\end{aligned} \tag{4.78}$$

Two quaternions are called *orthogonal* if $\mathbf{p}\bar{\mathbf{q}} \in \mathfrak{Vec}\mathbb{H}$. It is obvious that p,q are orthogonal if and only if there exists a pure quaternion $\mathbf{v}$ such that $\mathbf{p} = \mathbf{v}\mathbf{q}$. Quaternion multiplication preserves Euclidean norm, namely $\|\mathbf{pq}\| = \|\mathbf{p}\|\|\mathbf{q}\|$. If we identify $\mathbb{H}$ with $\mathbb{R}^4$, it is easy to check that $\mathbb{R} \times \{0\}$ is the center of $\mathbb{H}$. Therefore its complement is an invariant subspace under conjugation too. In fact, if we take a unit quaternion $\mathbf{q}$ and a pure quaternion $\mathbf{v}$, then

$$\begin{aligned}J(\cos\theta + \hat{n}\sin\theta)\mathbf{v} &:= (\cos\theta + \hat{n}\sin\theta)\mathbf{v}(\cos\theta - \hat{n}\sin\theta)\\ &= (\hat{n} \times \mathbf{v})\sin 2\theta + (1 - \cos 2\theta)[\hat{n}(\hat{n}\cdot\mathbf{v}) - \mathbf{v}] + \mathbf{v}\cos 2\theta\\ &\in \mathfrak{Vec}\mathbb{H}\end{aligned}$$

It is clear that J maps $\mathbb{B}_{\mathbb{H},1}$ onto $SO(3)$. In particular, when $\theta = 0$ or $\theta = \pi$, $J(\cos\theta + \hat{n}\sin\theta) = I$.

One can identify $\mathbb{H}$ with $\mathbb{C}^2$ if a quaternion is rewritten in the form $\mathbf{q} = q_0 + q_1\mathbf{i} + (q_2 + q_3\mathbf{i})\mathbf{j} = z + w\mathbf{j}$, and the multiplication rule here is $(z_1 + w_1\mathbf{j})(z_2 + w_2\mathbf{j}) = (z_1z_2 - w_1\bar{w}_2) + (z_1w_2 + w_1\bar{z}_2)\mathbf{j}$. It is easy to check that

$$z + wj \in \mathbb{B}_{\mathbb{H},1} \to U_{z,w} := \begin{pmatrix} z & w \\ -\bar{w} & \bar{z} \end{pmatrix} \in SU(2) \tag{4.79}$$

is an isomorphism. In other word $SU(2)$ is one-to-one parameterized by unit quaternions. The exponential map from the Lie algebra $su(2)$ has the explicit form

$$exp(\mathbf{A}) = \cos\theta\mathbf{I} + \frac{\sin\theta}{\theta}\mathbf{A} \quad \text{with } \mathbf{A} = \theta\begin{pmatrix} in_1 & n_2 + in_3 \\ -n_2 + in_3 & -in_1 \end{pmatrix}$$

In conclusion we have the identification $SO(3) = SU(2)/\pm\mathbf{I}$. Another way to look at the identification is through

$$\tilde{J} : g \in SU(2) \to [M \to gMg^{\dagger}] \in SO(3) \tag{4.80}$$

where M is a complex Hermitian matrix satisfying $TrM = 0$. Since $U(1) = \{g \in SU(2) : Ad_{\sigma_3}g = g\}$ and it is connected, $\tilde{J}(g)\sigma_3$ gives a map from $SU(2)$ into $\mathbb{S}^2$, where σ_3 is the Pauli matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, hence there is an identification between $SU(2)/U(1)$ and $\mathbb{S}^2$.

Six definitions of bi-invariant inner-product induced metrics on the rotation groups are compared in [56] and proved to be functional equivalent in the sense that there exist positive continuous strictly increasing functions h_i such that $h_i \circ \Phi_1 = \Phi_i$ with $i = 2, \cdots, 6$ while some of them are boundedly equivalent. These metrics in general can be classified into three types, one measures by using the quaternion difference, for instance

$$\Phi_1(\mathbf{p}, \mathbf{q}) = \arccos(|\mathbf{p} \cdot \mathbf{q}|)$$

another measures the derivation from the identity matrix in the Euclidean space, for instance

$$\Phi_2(\sigma_1, \sigma_2) = \|\mathbf{I} - \sigma_1\sigma_2^T\|_F$$

with $\|\cdot\|_F$ the Frobenius norm; the other measures in the Riemannian way, namely by the distance in Lie algebra

$$\Phi_3(\sigma_1, \sigma_2) = \|\log(\sigma_1\sigma_2^T)\|$$

with $\|\cdot\|$ being either the Frobenius norm or $\|S\|_H^2 = \frac{1}{2}\mathrm{tr}(S^TS)$ for $S \in so(3)$. Some characterization of mean rotation with respect to Φ_1 and Φ_2 is given in [39]. In particular there is

$$\underset{\sigma\in SO(3)}{\arg\min}\sum_{i=1}^{N}\Phi_2(\sigma, \sigma_i)^2 = \underset{\sigma\in SO(3)}{\arg\min}\Phi_2(\frac{1}{N}\sum_{i=1}^{N}\sigma_i, \sigma)$$

namely the projection of $\frac{1}{N}\sum_{i=1}^{N}\sigma_i$ onto the rotation group is the rotation mean with respect to Φ_2; and for $\sigma_1, \cdots, \sigma_N$ belonging to a same one-parameter subgroup of $SO(3)$ such that $\Phi_3(\sigma_i, \sigma_j) < \sqrt{2}\pi$ for any i, j with respect to Frobenius norm, there is an explicit expression

$$\arg\min_{\sigma \in SO(3)} \sum_{i=1}^{N} \Phi_3(\sigma, \sigma_i)^2 = \sigma_1(\sigma_1^T\sigma_2(\sigma_2^T\sigma_3(\cdots\sigma_{N-1}(\sigma_{N-1}^T\sigma_N)^{\frac{1}{2}})^{\frac{2}{3}}\cdots)^{\frac{N-2}{N-1}})^{\frac{N-1}{N}}$$

meanwhile

$$\sum_{i=1}^{N} \log(\sigma_i^T\sigma) = 0$$

is a necessary but not sufficient condition such that $\sum_{i=1}^{N}\Phi_3(\sigma_i, \sigma)^2$ achieves local minima.

Bibliography

[1] Barbara Zwicknagl, *Power series kernels*, Constr. Aporox. (2009)29:61-84

[2] Yanmu Zhou, *Arrangements of points on the sphere*, Thesis, University of South Florida, (1995)

[3] A. Zettl, *Sturm-Liouville theory*, AMS. (2005)

[4] L. Younes, *Shapes and diffeomorphisms*, Springer (2010)

[5] Yuan Xu, *Funk-Hecke formula for orthogonal polynomials on spheres and balls*, Bull. London Math. Soc. 32, 447-457, (2000)

[6] Yuan Xu and E.W. Cheney, *Strictly positive definite functions on spheres*, Proc. Amer. Math. Soc., Vol. 116, No. 4 (Dec., 1992), 977-981

[7] Kun Xu, Wei-Lun Sun, Zhao Dong, Dan-Yong Zhao, Run-Dong Wu, and Shi-Min Hu, *Anisotropic spherical Gaussians*, ACM Trans. Graphics, Vol. 32 Issue 6, November (2013)

[8] Guoliang Xu, *Discrete Laplace-Beltrami operator on sphere and optimal spherical triangulations*, Int. J. Comput. Geom. Appl. Vol.16, No.1 (2006) 75-93

[9] Y. Wiaux, J.D. McEven, P. Vandergheynst, and O. Blanc, *Exact reconstruction with directional wavelets on the sphere*, Mon. Not. R. Astron. Soc.000, 1-21, (2007)

[10] Kunyang Wang and Luoqing Li, *Harmonic analysis and approximation on the unit sphere*, Science Pr. (2000)

[11] N. Virchenko and I. Fedotova, *Generalized associated Legendre functions and their applications*, World Scientific Publishing (2001)

[12] Grant V. Welland, *Beyond wavelets*, Academic Press (2003)

[13] G.N. Watson, *A treatise on the theory of Bessel functions*, 2nd Edition, Cambridge University Press (1944)

[14] J. Wang and W. Huang, *Image segmentation with eigenfunctions of an anisotropic diffusion operator*, IEEE T. Image. Process. (2016)

[15] M.P. Wand and M.C.Jones, *Kernel smoothing*, first edition (1995)

[16] N.Ja. Vilenkin and A.U. Klimyk, *Representation of Lie groups and special functions*, Vol.1 and Vol.2, Mathematics and its applications(Soviet series) (1991-1993)

[17] P. Vandergheynst, J-P. Antoine, L. Jacques, and M. Morvidone, *Stereographic wavelet frames on the sphere*, Appl. Comput. Harmon. Anal. vol.19, num.2, p.223-252, (2005)

[18] G. Szegő, *Orthogonal polynomials*, AMSCP, Vol.23 (1975)

[19] W. Sweldens, *The lifting scheme: a construction of second generation wavelets*, SIAM J. Math Anal. Vol.29, No.2, p.511-546, (1998)

[20] Yizhi Sun and Zhilin Sun, *Generating probability distributions on intervals and spheres with application to finite element method*, submitted

[21] E.M. Stein, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press (1971)

[22] J.-L. Starck, Y. Moudden, P. Abrial, and M. Nguyen, *Wavelets, ridgelets and curvelets on the sphere*, Astron. Astrophys. 446, p.1191–1204 (2006)

[23] J.-L. Starck, F. Murtagh, and J.M Fadili, *Sparse image and signal processing: wavelets, curvelets, morphological diversity*, Cambridge University Press (2010)

[24] G.F. Smoot, *Nobel Lecture: Cosmic microwave background radiation anisotropies: their discovery and utilization*, Rev. Mod. Phys., Vol. 79, No. 4, Oct.–Dec. (2007)

[25] P. Schröder and W. Sweldens, *Spherical wavelets: efficiently representing functions on the sphere*, in the book "Wavelets in the geosciences", p158-188, Springer (2000)

[26] I.J. Schoenberg, *Positive definite functions on spheres*, Duke Math J. vol.9, No.1, p.96-108 (1942)

[27] H. Harbrecht and R. Schneider, *Rapid solution of boundary integral equaitons by wavelet Galerkin schemes*, in the book "Multiscale, nonlinear and adaptive approximation", Springer (2009)

[28] D. Roşca, *Locally supported rational spline wavelets on a sphere*, Math Comput. Vol. 74, No. 252, p.1803–1829 (2005)

[29] F. Narcowich, P. Petrushev, and J. Ward, *Decomposition of Besov and Triebel-Lizorkin spaces on the sphere*, J. Funct. Anal. 238, 530-564 (2006)

[30] F. Narcowich, P. Petrushev, and J. Ward, *Localized tight frames on spheres*, SIAM J. Math. Anal. 38 p.574-594 (2006)

[31] G. Plonka and D. Roșca, *Easy path wavelet transform on triangulations of the sphere*, Math. Geosci. (2010) 42:839-855

[32] A.A. Penzias and R.W. Wilson, *A measurement of excess antenna temperatures at 4080 Mc/s*, Astrophys. J. Vol.142(1965), p.419-421

[33] S. Pawelke, *Über die Approximationsordnung bei Kugelfunktionen und algebraischen Polynomen*, Tôhoku Math. Journ. 24 p.473-486 (1972)

[34] Roland Opfer, *Multiscale kernels*, Adv. Comput. Math. 25: 357-380 (2006)

[35] Roland Opfer, *Tight frame expansions of multiscale reproducing kernels in Sobolev spaces*, Appl. Comput. Harmon. Anal. 20 (2006) 357-374

[36] C. Müller, *Analysis of spherical symmetries in Euclidean spaces*, Springer (1997)

[37] J.L. Müller and S. Siltanen, *Linear and nonlinear inverse problems with practical applications*, SIAM (2012)

[38] M.J. Mohlenkamp, *A fast transform for spherical harmonics*, J. Fourier. Anal. Appl. Vol.5(2/3), p.159-184 (1999)

[39] Maher Moakher, Means and averaging in the group of rotations, SIAM J. Matrix Anal. Appl. Vol.24, No.1, p.1-16 (2002)

[40] V. Michel, *Lectures on constructive approximation*, Springer (2013)

[41] J.D. McEwen, C. Durastanti, and Y. Wiaux, *Localisation of directional scale-discretised wavelets on the sphere*, Appl. Comput. Harmon. Anal. 44 (2018) 59-88

[42] J.D. McEwen, P. Viela, Y. Wiaux, R.B. Barreiro, L. Cayón, M.P.Hobson, A.N. Lasenby, Martínez-González, and J.L. Sanz, *Cosmological applications of a wavelet analysis on the sphere*, J. Fourier Anal. Appl. Vol.13, Issue 4 (2007)

[43] S. Mallat, *A wavelet tour of signal processing: the sparse way*, third edition (2009)

[44] Jianwei Ma and Gerlind Plonka, *The curvelet transform-a reveiw of recent applications*, IEEE Signal Processing Magazine 118 (2010)

[45] Fangyang Lu and Hongwei Sun, *Positive definite dot product kernels in learning theory*, Adv. Comput. Math. 22:181-198 (2005)

[46] Christian Lessig and Eugene Fiume, *SOHO: orthogonal and symmetric Haar wavelets on the sphere*, ACM Trans. on Graphics, Vol. 27, No. 1 (2008)

[47] G. Kutyniok and D. Labate, *Resolution of the wavefront set using continuous shearlets*. Trans. Amer. Math. Soc. 361(5), p.2719–2754, (2009)

[48] A. Kunoth and J. Sahner, *Wavelets on manifolds: an optimized construction*, Math. Comput. Vol.75, No.255, p.1319-1349 (2006)

[49] J. Korevaar, *Tauberian theory, a century of developments*, Springer (2004)

[50] Matthias Keller, *Geometric and spectral consequences of curvature bounds on tessellations*, in the book "Modern approaches to discrete curvature" (2017)

[51] I. Iglewska-Nowak and M. Holschneider, *Frames of Poisson wavelets on the sphere*, Appl. Comput. Harmon. Anal. 28(2010) 227-248

[52] I. Iglewska-Nowak, *Frames of directional wavelets on n-dimensional spheres*, Appl. Comput. Harmon. Anal. 43(2017) 148–161

[53] I. Iglewska-Nowak, *Semi-continuous and discrete wavelet frames on n-dimensional spheres*, Appl. Comput. Harmon. Anal. 40(2016) 529–552

[54] I. Iglewska-Nowak, *Continuous wavelet transforms on n-dimensional spheres*, Appl. Comput. Harmon. Anal. 39(2015) 248–276

[55] S. Hubbert and B. Baxter, *Radial basis functions for the sphere*, London: Birkbeck ePrints (2001)

[56] Du Q. Huynh, *Metrics for 3D rotations: comparison and analysis*, J. Math. Imaging Vis. (2009) 35: 155-164

[57] L. K. Hua, *Harmonic analysis of functions of several complex variables in the classical domains*, Science Press, Peking (1958)

[58] Wayne Hu and Scott Dodelson, *Cosmic microwave background anisotropies*, Annu. Rev. Astron. and Astrophys. (2002)

[59] T. Hou and H. Qin, *Continuous and discrete Mexican hat wavelet transforms on manifolds*, Graph. Models 74, p.221-232 (2012)

[60] E. Hilb, *Über die Laplacesche Reihe*, Math. Z. vol.5 (1919), p.17-25; vol.8 (1920), p.79-90

[61] Yu. Higuchi, *Combinatorial curvature for planar graphs*, J. Graph Theory 38:220-229 (2001)

[62] M. Gräf and D. Potts, *Sampling sets and quadrature formulae on the rotation group*, Numer. Func. Anal. Opt. 30 7-8 p.665-688 (2009)

[63] D. Geller and A. Mayeli, *Continuous wavelets on compact manifolds*, Math. Z. (2009) 262:895–927

[64] D. Geller and D. Marinucci, *Mixed needlets*, J. Math. Anal. Appl.375 p.610-630 (2011)

[65] I.M. Gel'fand, R. A. Minlos and Z. Ja Šapiro, *Representations of the rotation group and of the Lorentz group, and their applications*, Fizmatgiz, Moscow, 1958; English translat., Macmillan, New York.

[66] W. Freeden and U. Windheuser, *Spherical wavelet transfrom and its discretization*, Adv. Comput. Math. 5(1996) 51-94

[67] W. Freeden, T. Gervens, M. Schreiner, *Constructive approximation on the sphere with applications to geomathematics*, Oxford University Press (1998)

[68] W. Freeden and M. Gutting, *Special functions of mathematical (geo-)physics*, Springer (2013)

[69] W.T. Freeman and E.H. Adelson, *The design and use of steerable functions*, IEEE Trans. Patt. Anal. and Machine Intell. Vol. 13, No. 9, p.891-906, (1991)

[70] M.S. Floater and E.G. Quak, *Linear independence and stability of piecewise linear prewavelets on arbitrary triangulations*, SIAM J. Numer. Anal., Math. No. 1, p.58-79 (2000)

[71] A.R. Edmonds, *Angular momentum in quantum mechanics*, Princeton University Press (1957)

[72] Tarn Duong, *Bandwidth selectors for multivariate kernel density estimation*, doctor thesis, (2005)

[73] Javier Duoandikoetxea, *Fourier analysis*, AMS, (2000) translated and revised by David Gruz-Uribe.

[74] S. Dahlke, W. Dahmen, I. Weinreich and E. Schmitt, Multiresolution analysis and wavelets on $\mathbb{S}^2$ and $\mathbb{S}^3$, Numer. Func. Anal. Opt. 16:1-2, 19-41 (1995)

[75] Josef Dodziuk, *Difference equations, isoperimetric inequality and transience of certain random walks*, Trans. Amer. Math. Soc. Vol.284, No.2 (1984)

[76] Josef Dodziuk and W.S. Kendall, *Combinatorial Laplacians and isoperimetric inequality*, (1986)

[77] Minh N. Do and Martin Vetterli, *The contourlet transform: an efficient directional multiresolution image representation*, IEEE Trans Image Process. (2005) Dec;14(12):2091-106.

[78] Feng Dai and Yuan Xu, Approximation Theory and Harmonic Analysis on Spheres and Balls, Springer (2013)

[79] W. Dahmen and R. Schneider, *Wavelets on manifolds I: construction and domain decomposition*, SIAM J. Math. Anal. Vol.31, No.1, p.184-230, (1999)

[80] V. Climenhaga and A. Katok, *From groups to geometry and back*, AMS. (2017)

[81] O. Christensen, *An introduction to frames and Riesz bases*, second edition. Springer (2016)

[82] L. Cayon, J.L. Sanz, E.Martinez-Gonzalez, A.J.Banday, F. Argueso, J.E. Gallegos, K.M. Gorski and G.Hinshaw, *Spherical Mexican hat wavelet: an application to detect non-Gaussianity in the COBE-DMR maps*, Mon. Not. R. Astron. Soc. (2001)

[83] M. Calixto, J. Guerrero, and D. Roşca, *Wavelet transform on the torus: a group theoretical approach*, Appl. Comput. Harmon. Anal. 38 (2015) 32-49

[84] Xiaohao Cai, Christopher G. R. Wallis, Jennifer Y. H. Chan, and J.D. McEwen, *Wavelet-based segmentation on the sphere*, arXiv:1609.06500v1 [cs.CV] 21 Sep 2016

[85] Thomas Bülow, *Multiscale image processing on the sphere*, Pattern Recognition, DAGM2002, p.609-617 (2002)

[86] I. Bogdanova, P. Vandergheynst, and J-P. Antoine, *Continuous wavelet transform on the hyperboloid*, Appl. Comput. Harmon. Anal. 23 (2007) p.285–306

[87] J.J. Benedetto and M. Fickus, *Finite normalized tight frames*, Adv. Comput. Math. 18: 357-385 (2003)

[88] Frank Bauer, Bobo Hua, Jürgen Jost, Shiping Liu, and Guofang Wang, *The Geometric meaning of curvature: local and nonlocal aspects of Ricci curvature*, in the book "Modern approaches to discrete curvature" (2017)

[89] P. Auscher, G. Weiss, and M. Victor, *Local sine and cosine bases of Coifman and Meyer and the construction of smooth wavelets*, Wavelets-A Tutorial, Academic Press, p.237-256, (1991)

[90] P. Audet, *Directional wavelet analysis on the sphere: application to gravity and topography of the terrestrial planets*, J. Geophys. Res., 116, E01003 (2011)

[91] N. Aronszajn, *Theory of reproducing kernels*, Amer. Math. Soc. Trans. 63, p.337-404 (1950)

[92] J-P. Antoine, L. Demanet, L. Jacques, and P. Vandergheynst, *Wavelets on the sphere: implementation and approximations*, Appl. Comput. Harmon. Anal. vol.13, p.177-200 (2002)

[93] J-P. Antoine and P. Vandergheynst, *Wavelets on the 2-sphere: a group-theoretical approach*, Appl. Comput. Harmon. Anal. vol.7 (1999) p.262-291

[94] J-P Antoine, D. Roşca, and P. Vandergheynst, *Wavelet transform on manifolds: old and new approaches*, Appl. Comput. Harmon. Anal. 28 (2010) p.189–202

[95] M. Andreux, E. Rodola, M. Aubry, and D. Cremers, *Anisotropic Laplace-Beltrami operators for shape analysis*, Computer Vision - ECCV 2014 Workshops. vol.8928, p.299-312 (2015)

[96] W.O. Amrein, A.M. Hinz, and D.B. Pearson, *Sturm-Liouville theory: past and present*, Springer (2005)

[97] W.K. Allard, G. Chen, and M. Maggioni, *Multi-scale geometric methods for data sets II: geometric multi-resolution analysis*, Appl. Comput. Harmon. Anal. 32 p.435-462 (2012)

[98] M.A. AL-Gwaiz, *Sturm-Liouville theory and its applications*, Springer (2007)