Republic of Iraq
Ministry of Higher Education
and Scientific Research
University of Babylon
College of Education for pure
Sciences
Department of Mathematics

Functions Approximation by Spectral Graph Wavelets

Research

Submitted to College of Education for Pure
Sciences of University of Babylon as Partial Fulfillment of the
Requirement for the Degree of
Higher Diploma Education / Mathematics

by

Ali Saeed Mnahi Agel ALmshlab

Supervised by

Ass.Prof.Dr. Hawraa Abbas Fadhil Almurieb

2022

بسم الله الرحمن الرحيم

﴿ وَأَنْزَلَ اللهُ عَلَيْكَ الكِتَابَ وَالحِكْمَةَ

وَعَلَّمَكَ مَا لَمْ تَكُنْ تَعْلَمُ

وَكَانَ فَضْلُ اللهِ عَلَيْكَ عَظِيمًا ﴾

النساء ١١٣

Dedication

To the one who showered us
with her love all her life...
To dear mother... May God
prolong her life in obedience to
him...

To Every One We Love
And Every One Who Care
About Us

ACKNOWLEDGMENTS

After the completion of the work, there is nothing more beautiful or sweeter than praise and thanks be to God, as it should be considered as the majesty of his countenance, the greatness of his authority, and the best for his great bounty and great benevolence for what he has bestowed upon me to complete this humble research.

I would like to express my sincere thanks and gratitude to my supervisor Dr. Hawraa Abbas Fadhil for the continuous support. She is always motivating, inspiring, encouraging and her guidance helped me in all the time of the research .

ABTRACT

Approximating function by using spectral graph wavelets is an interesting direction in approximation theory. It opens wide doors to a new world of function approximation. We essential to well choosing the space of the functions that are approximated by spectral graph wavelets. L_p spaces of functions are fantastic choices to study It is more interesting to take the value $0 < p < 1$. In this research , new formulas of spectral graph wavelets were constructed and proved to get good rates of approximation fundamental properties of L_p graph wavelets transform s (L_p GWT) and studied . Inversion , scaling limit and approximation wavelets. finally , existence of best approximation can be concluded here for graph functions.

CONTENTS

FIGURES

SYMBOLS

Symbol	Definition	page
L_p	Lebesgue space	6
$E_n^*(.)$	Degree of best approximation	11
$\omega_f\ (\delta)$	Modulus of continuity	13
$\mathcal{L}^{norm}$	Normalized Laplacian	19
ψ	Mother wavelet	21
ε	Epsilon	28
$\|.\|_p$	L_p_norm	30

ABBREVIATION

Symbol	Definition
CWT	**Continuous Wavelet Transform**
L_p SGWT	L_p Spectral Graph Wavelet Transform

PUBLICATION

Functions Approximation by Spectral Graph Wavelets

Submitted to

Al-Bahir Joumal for Engineering and Pure Sciences

INTRODUCTION

Wavelets are mathematical functions that cut up data into different frequency components, and then study each component with a resolution matched to its scale. Many interesting scientific problems involve analyzing and manipulating structured data. Such data often consist of sampled real-valued functions defined on domain sets themselves having some structure. The simplest such examples can be described by scalar functions on regular Euclidean spaces, such as time series data, images or videos. However, many interesting applications involve data defined on topologically complicated domains. Examples include data defined on network-like structures, data defined on manifolds or irregularly shaped domains. Many signal processing techniques are based on transform methods, where the input data is represented in a new basis before analysis or processing. One of the most successful types of transforms in use is wavelet analysis. Wavelets have proved over the past 25 years to be an exceptionally useful tool for signal processing. Much of the power of wavelet methods comes from their ability to simultaneously localize signal content in both space and frequency. For signals whose primary information content lies in localized singularities, such as step discontinuities in time series signals or edges in images, wavelets can provide a much more compact representation than either the original domain or a transform with global basis elements such as the Fourier transform.Classical wavelets are constructed by translating and scaling a single "mother" wavelet. The transform coefficients are then given by the inner products of the input function with these translated and scaled waveforms.This problem is working in the spectral graph domain[1], i.e. using the basis consisting of the eigenfunctions of the graph Laplacian $\mathcal{L}$.In this research, we introduce a new version of spectral graph wavelet transform in L_p space and study several of its properties. We also show that in the fine scale limit, for sufficiently regular g, the wavelets exhibit good localization properties.Numerous authors have introduced extensions and related transforms for signals on the plane and higher-dimensional spaces. By taking separable products of one-dimensional wavelets, one can construct orthogonal families of wavelets in any dimension [2]. However, this yields wavelets with often undesirable bias for coordinate axis directions. Additionally, this approach yields a number of wavelets that is exponential in the dimension of the space, and is unsuitable for data embedded in spaces of large dimensionality. Wavelet transforms have also been defined for certain non-Euclidean manifolds, most notably the sphere [3,4] and other conic sections [5][6]. Previous authors have explored wavelet transforms on graphs, albeit via different approaches to those employed in this paper. Crovella and Kolaczyk [7] defined wavelets on unweighted graphs for analyzing computer network traffic. Smalter et al. [8] used the graph wavelets of Crovella and Kolaczyk as part of a larger method for measuring structural differences between graphs representing chemical structures, for machine learning of chemical activities for virtual drug screening.Jansen et al. [9] develop a multiscale scheme for data on graphs based on lifting. Their scheme requires distances to be assigned to each edge, which for their examples are inferred from Euclidean distances when the graph vertices correspond to irregularly sampled points of Euclidean space. Murtagh [10] developed a Haar wavelet transform for rooted binary trees, known as dendrograms. This concept was expanded upon by Lee et al. [11] who developed the treelet transform, incorporating automatic construction of hierarchical trees for multivariate data.Our contribution to this research is through three chapters. Chapter one includes an introduction to the subject, basic definitions and paving theories for the birth of approximation of functions using the waves of the spectral statement. In addition, we stated the applied origin of the problem as well as the solutions proposed by the important theories in approximating continuous functions such as the theorems of existence and uniqueness. Then, we touch on graph theory because of its importance in our research. In chapter two, we defined L_p Spectral Graph Wavelet Transforms and study their important characteristics and theorems, to reach finally to the purpose of the research, that is to approximate functions of the space L_p, especially those functions of vertices in chapter three.

Chapter One

Introduction and Preliminaries

This present research research of function approximation by wavelets of spectral graph type needs some preliminaries concluding birth of the topic, main concepts, theorems, and previous results. In this chapter, we present the literature review concerning with our work in three hubs, the first is about approximation theory, the second is about spectral graphs, and the final is about wavelets. Finally, we link the axes together to focus on the later results.

1.1. Introduction to Approximation Theory

In 1853, the great Russian mathematician, P.L. Chebyshev, while working on a problem of linkages, devices which translate the linear motion of a steam engine into the circular motion of a wheel, considered the following problem in its primitive forms:

"Given a continuous function f defined on a closed interval $[a,b]$ and a positive integer n, can we "represent" f by a polynomial $p(x) = \sum_{k=0}^{n} ax^k$ of degree at most n, in such a way that the maximum error at any point x in $[a,b]$ is controlled. In particular, is it possible to construct p in such a way that the error $\max_{a\leq x\leq b}|f(x)-p(x)|$is minimized"?[12]

Many questions arose from this problem; Is this construction possible? How and When ? Is it always unique? Do the continuous functions only have approximation? What about other spaces? Are Polynomials the only approximations to all functions?

Actually, some of those questions have got answers while others are still being asked. To have a general look about the former results here, we need to present some notations concerning with this topic,

Definition 1.1.1.[12]

A normed space is a real vector space X ,equipped with a non-negative real valued function $\|\cdot\|$, that is for any vectors $x, y \in X$ and scalar $\alpha \in R$, we have

1. $\|x\| \geq 0, \ \ and \ \ \|x\| = 0 \leftrightarrow x = 0$
2. $\|\alpha x\| = |\alpha|\|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$

Any norm on X induces a metric or distance function by setting

$$\text{dist}(x, y) = \| x - y\|.$$

Definition 1.1.2.[13]

The L_p space for $0 < p < \infty$ is given by

$$L_p([a, b]) = \left\{f: [a, b] \to \mathbb{R}, \middle| \, f \text{ measurable and } \|f\|_p < \infty\right\},$$

where

$$\|f\|_p = \left\{\sum_{i=1}^{N} |f(x_i)|^p\right\}^{1/p},$$

Note that L_p is norm for the case $p > 1$, but it is not for the case of $p < 1$, so we give the definition of Quasi-Norm depending on the following relation,

Theorem 1.1.3.(Quasi-Triangle inequality)[14]

Let $0 < p < 1$ then

$$\|x + y\|_p \leq C\left(\|x\|_p + \|y\|_p\right),$$

C is anon_negative constant

Definition 1.1.4.(Quasi-normed Space)[14]

A quasi-norm on a real vector space X is a real- valued function $\|\cdot\| : X \rightarrow [0, \infty)$ such that

(1) $\| x\| = 0$ if and only if $x = 0$,

(2) $\|\alpha x\| = |\alpha|\| x\|$, $\alpha \in R$, $x \in X$,

(3) $\|x + y\| \leq C(\|x\| + \|y\|)$, $x, y \in X$,

Where $C \geq 1$ is a constant independent of $x, y \in X$. The smallest possible constant $C = C_X \geq 1$ is called the quasi-triangle constant of $X = (X, \| . \|)$

Theorem 1.1.5.[13]

Let $1 < p < \infty$, and let $1 < q < \infty$ be defined by

$\frac{1}{p} + \frac{1}{q} = 1$ that is $q = \frac{p}{p-1}$ then for any $a, b \geq 0$ we have

1.Young's Inequality

$$ab \leq \frac{1}{p} a^p + \frac{1}{q} b^q$$

Moreover, equality can only occur if $a^p = b^q$. We refer to p and q as conjugate exponents; note that p satisfies $p = \frac{q}{q-1}$. Please note that the case $p = q = 2$ yields the familiar arithmetic-geometric mean inequality.

2.Holder's Inequality

$$\|fg\|_1 \leq \|f\|_p \|g\|_q \quad \text{when} \quad \frac{1}{p} + \frac{1}{q} = 1$$

the case $p = q = 2$ yields the familiar Cauchy-Schwarz inequality.

3. Minkowski's Inequality

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p$$

The properties in the above theorem are not satisfied for the case $0 < p < 1$, but equivalently we have

Theorem 1.1.6.[13]

Let $0 < p < 1$, we have

1. **Quasi-Triangle Inequality**

$$\|f + g\|_p \leq 2^{\frac{1}{p}-1}\left(\|f\|_p + \|g\|_p\right)$$

2. **Minkowski's Inequality**

$$\|f + g\|_p \geq \|f\|_p + \|g\|_p$$

We are able now to introduce the abstract version of the problem of approximation as follow,

Definition 1.1.7.[15]

Let Y be a nonempty subset of the space X and let $x \in X$. An element $y_o \in Y$ is called **a best approximation**, or nearest point, to x from Y if

$$\|x - y_o\| = d(x, Y),$$

where $d(x, Y) \coloneqq \inf_{y \in Y} \|x - y\|$. The number $d(x, Y)$ is called the distance from x to Y, or the error in approximating x by Y. The (possibly empty) set of all best approximations from x to Y is denoted by P_x. Thus

$$P_x := \{y \in Y : \|x - y\| = d(x, Y)\}.$$

Existence Theorem 1.1.8.[12]

Let Y be a finite-dimensional subspace of a normed linear space X, and let $x \in X$. Then, there exists a (not necessarily unique) $y^* \in Y$ such that

$$\|x - y^*\| = \min_{y \in Y} \|x - y\|$$

for all y ∈ Y . That is, there is a best approximation to x by elements of Y One of the earliest cases of existence of approximation is the following one that was originally discovered by Weierstrass.

WEIERSTRASS Theorem 1.1.9.[12]

For each $f \in C[\,a,b\,]$, and each positive integer n, there is a (not necessarily unique) polynomial $p_n^* \in P_n$ such that

$$\|f - p_n^*\| = \min_{p \in P_n} \|f - p\|$$

The set Y_x has properties given in the following theorem

Theorem 1.1.10.[12]

Let Y be a subspace of a normed linear space X, and let $x \in X$. The set Y_x, consisting of all best approximations to x out of Y , is a bounded, convex set.

Definition 1.1.11.[15]

Let $S \in R^n$. If the line segment between any two points in S liesin S, i.e.,

$$\lambda x_1 + (1-\lambda)x_2 \in S \quad \text{all} \quad \lambda \in [0\,,1]$$

then S is said to be **convex.**

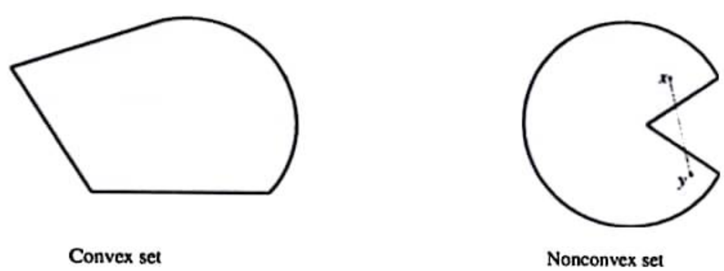

Figure 1.1 Convex set and Nonconvex set

Examples 1.1.12.[15]

1.Any closed ball centered at x with radius r, which is defined by

$$B[x,r] := \{y \in X \;\; \|x - y\| \le r\},$$

is convex.

2.Any open ball centered at x with radius r , which is defined by

$$B(x,r) := \{y \in X \ \ \|x - y\| < r\},$$

is convex.

Definition 1.1.13.[15]

We say that S strictly convex normed linear space if

$$f_1 \neq f_2 \ , \qquad \|f_1\| = r \ , \qquad \|f_2\| = r$$

then $\|\lambda f_1 + (1-\lambda) f_2\| < r$ for all λ satisfying $0 < \lambda < 1$

Uniqueness Theorem 1.1.14.[12]

X has a strictly convex norm if and only if the triangle inequality is strict on non-parallel vectors; that is, if and only if

$x \neq \alpha y, y \neq \alpha x$, for all $\alpha \in R$ Then $\|x + y\| < \|x\| + \|y\|$.

Corollary 1.1.15.[12]

If X has a strictly convex norm, then, for any subspace Y of X and any point $x \in X$, there can be at most one best approximation to x out of Y . That is, Y_x is either empty or consists of a single point.

Definition 1.1.16.[12]

The degree of best approximation of the function $f \in X$ is given by

$$E_n^*(f) = E_n(f) = \inf_{\emptyset \in Y} \|f - \emptyset\|$$

where $\emptyset \in Y$ is a function that is a best approximation of f out of Y.

Example 1.1.17.

consider $X = R^2$ under the norms

i. $\|(x_1, x_2)\|_\infty = max\{|x_1|, |x_2|\}$,

ii. $\|(x_1, x_2)\|_2 = \sqrt{(x_1)^2 + (x_2)^2}$,

iii. $\|(x_1, x_2)\|_1 = |x_1| + |x_2|$

Approximation of (2,0) from $Y = \{(1, y)\}$ is the set $\{(1, y) : \ -1 \leq y \leq 1\}$

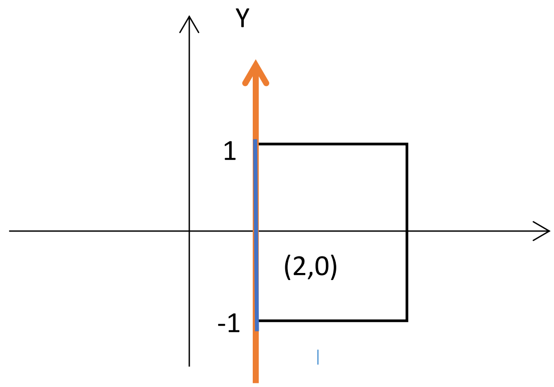

Figure 1.2 Example in L_∞

Approximation of (2,0) from $Y = \{(1, y)\}$ is (1,0)

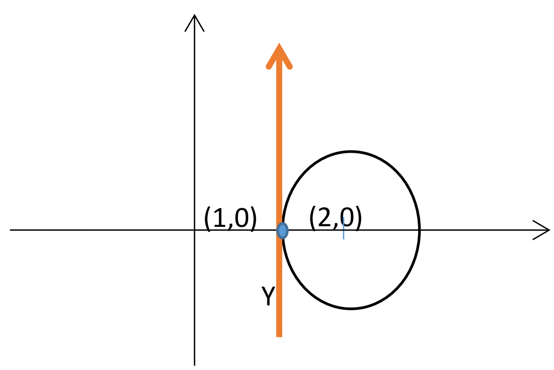

Figure 1.3 Example in L_2

Approximation of (2,0) from $Y = \{(1, y)\}$ is (1,0)

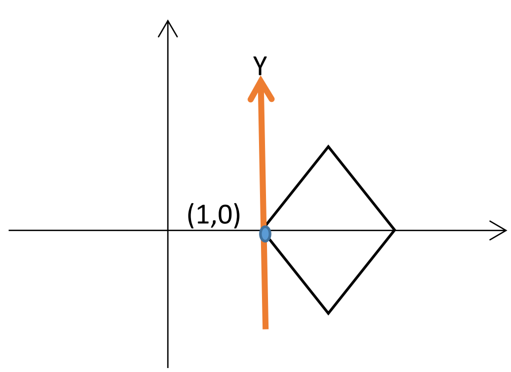

Figure 1.4 Example in L_1

1.1.18. Example for $C[a,b]$ [12]

Define $C[\,a,b\,]$ to be the set of continuous functions on $[a,b]$. We're interested to get approximations to elements of $C[\,a,b\,]$ by elements of $Y = \mathrm{p_n}$, the subspace of all polynomials of degree at most n in $C[\,a,b\,]$. $\mathrm{p_n}$ is a finite-dimensional subspace of $C[\,a,b\,]$ of dimension exactly $n+1$.

$$\mathrm{Dim}(\mathrm{p_n}) = n+1 = \mathrm{Dim}\left(R^{(n+1)}\right)$$

Since there is an isometry between $\mathrm{p_n}$ and $R^{(n+1)}$ $\left(\mathrm{p_n} \cong R^{(n+1)}\right)$

It follows from the Weierstrass theorem, for example, that each $f \in C[\,a,b\,]$ has distance 0 from p but, since not every $f \in C[\,a,b\,]$ is a polynomial, for example, the function $f(x) = x\, sin(1/x)$ is continuous on $[\,0,1\,]$ but can't possibly agree with any polynomial on $[\,0,1\,]$.

Since $xsin\left(\frac{1}{x}\right) = x\left(\frac{1}{x} - \frac{1}{x^3} + \frac{1}{x^5} - \cdots\right) = \left(1 - \frac{1}{x^2} + \frac{1}{x^4} - \cdots\right)$

It does not represent a polynomial.

The key to the problem of polynomial approximation is the fact that each $\mathrm{p_n}$ is finite-dimensional. To see this, it will be most efficient to consider the abstract setting of finite-dimensional subspaces of arbitrary normed spaces .

Definition 1.1.19.[16]

We say that $\{f_n\}$ is uniformly bounded on E i f there exists a number M such that

$$|f_n(x)| < M \qquad (x \in E\,, n = 1,2,3,\dots).$$

Sometime, one prefer to improne the degree of approximation .

With the following modeley of continuity . It is a measure of ε in the uniform continuity definition,

1.1.20. Modulus of Continuity [12]

The modulus of continuity of a bounded function f on the interval $[\,a, b\,]$ is defined by

$$\boldsymbol{\omega_f\,(\delta) = \omega_f\,([\,a, b\,];\,\delta) = sup\{|f(x)\,-\,f(y)|\,:\,x, y\,\in\,[\,a, b\,], |x\,-\,y|\,\leq\,\delta\}}$$

for any $\delta > 0$.

In the following section, we present some major concepts about graphs, and spectral graphs as well.

1.2. Spectral Graphs

Definition 1.2.1.[17]

A graph $G = (V, E)$ is a pair (V, E) consisting of a finite non- empty set V of vertices (called also points, nodes, or just dots), and a finite set E of edges, where some pair of vertices is connected by one or more edges.

Example 1.2.2.

Here we have a graph G_1 with vertices $V = \{a, b, c, d, e, f\}$ and the edges $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8, e_9, e_{10}, e_{11}, e_{12}\}$

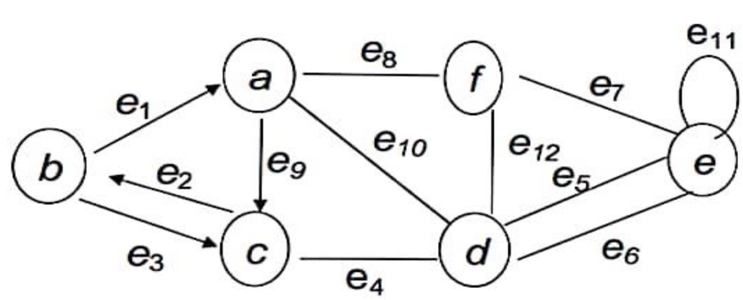

Figure 1.5. Graph G_1

Definition 1.2.3.[18]

A path is a sequence of edges which joins a sequence of vertices which are all distinct. In graph G_1 we have many different paths but we only gave two examples of paths:

$$e_1\ e_{10} e_5\ e_7\ ,\ \ e_1 e_8\ e_7\ e_6 e_4$$

Definition 1.2.4.[18]

An undirected graph is a graph in which any edges in the graph have no specific direction. While a directed graph is a graph which any edge in the graph have a specific direction.

Definition 1.2.5.[18]

A mixed graph is a graph which combines both undirected and directed properties. Graph in Figure 1.6. is example for mixed graph.

Definition 1.2.6.[19]

Two vertices (v_1, v_2) are adjacent (neighbours) if they are joined by the same edge e_1 and the edge e_1 is said to be incident with the vertices (v_1, v_2).

If we take graph G_1 as example the vertex a in graph G_1 is neighbour of four vertices b, c, d, f and b is only neighbour of two vertices a, c.

Definition 1.2.7.[18]

A weighted graph $G = \{E, V, w\}$ consists of a set of vertices V , a set of edges E, and a weight function $w : E \to R^+$ which assigns a positive weight to each edge. We consider here only finite graphs where $|V| = N < \infty$.

Definition 1.2.8.[19]

The degree of a vertex in unweighted graph is the number of edges that are incident with this vertex. For example, vertex a in graph G_1 has degree four and b has degree three.

On the other hand, the degree of a vertex m in a weighted graph, $d(m)$, is the sum of weights of all the edges incident to it. i.e.

$$d(m) = \sum_{n} a_{m,n}$$

Definition 1.2.9.[18]

A loop is an edge which joins one vertex to itself. Example: in graph G_1 we have only one loop namely the edge e_{11}.

Loops implies nonzero diagonal elements in the adjacency matrix, (defined later)

Definition 1.2.10.[18]

Parallel edges are two or more different edges which join the same pair of vertices. Example: In the Graph G_1 only two pair edges are parallel: (e_5, e_6)

Definition 1.2.11.[18]

A simple graph is a graph which contains no loops and no parallel edges.

Definition 1.2.12.[19]

A graph is called k-regular if the degree of every vertex is k.

Definition 1.2.13.[19]

We abbreviate $d(v_i)$ as d_i. The degree matrix $D(G)$, is the diagonal matrix, $D(G) = \text{diag}(d_1, \ldots, d_m)$.

Example 1.2.14.

Here we have unweighted graph G_2 (V,E) where $V = \{v_1, v_2, v_3, v_4, v_5\}$ and $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$

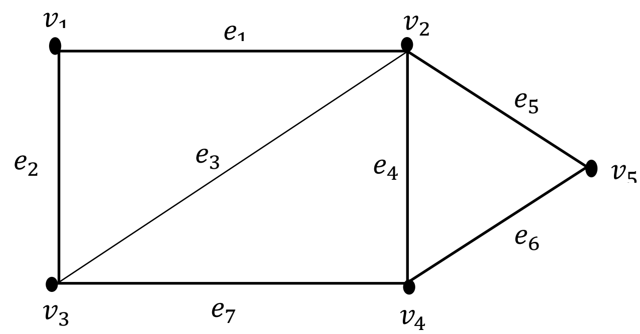

Figure 1.7. Graph G_2

For example, for graph G_2, we have

$$D(G_2) = \begin{pmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix}$$

Unless confusion arises, we write D instead of $D(G)$.

Now , consider $w: E \to \{1,3,4,5\}$ whene

$w(e_1) = w(e_3) = 1$

$w(e_2) = w(e_4) = 3$

$w(e_5) = w(e_6) = 4$

$w(e_7) = 5$, then

$$D(G_2) = \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 9 & 0 & 0 & 0 \\ 0 & 0 & 9 & 0 & 0 \\ 0 & 0 & 0 & 12 & 0 \\ 0 & 0 & 0 & 0 & 8 \end{pmatrix}$$

Definition 1.2.15.[19]

A simple unweighted graph G with vertex set $V(G) = \{ v_1 , \ldots , v_n\}$ and edge set $E(G) = \{e_1, \ldots , e_n\}$ can also be described by means of matrices. One such matrix is the $n \times n$ adjacency matrix

$$A(G) = [a_{ij}] \text{ , where } a_{ij} = \begin{cases} 1 , & \text{if } e \in E(G) \\ 0 , & \text{if } e \notin E(G) \end{cases} \text{ , where } e = v_i v_j$$

Here is the adjacency matrix of a graph G_2

$$A = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix}$$

Definition 1.2.16.[19]

The adjacency matrix A for a weighted graph G is the $N \times N$ matrix with entries $a_{m,n}$ where

$$a_{m,n} = \begin{cases} w(e) & if \;\; e \in E \text{ connects vertices } m \text{ and } n \\ 0 & \text{otherwise} \end{cases}$$

$$A = \begin{pmatrix} 0 & 1 & 3 & 0 & 0 \\ 1 & 0 & 1 & 3 & 4 \\ 3 & 1 & 0 & 5 & 0 \\ 0 & 3 & 5 & 0 & 4 \\ 0 & 4 & 0 & 4 & 0 \end{pmatrix}$$

Definition 1.2.17.[20]

Every real-valued function $\boldsymbol{f} : V \to R^N$ on the vertices of the graph G can be viewed as a vector in R^N , where the value of f on each vertex defines each coordinate. This implies an implicit numbering of the vertices. We adopt this identification , and will write $f \in R^N$ for functions on the vertices of the graph, and $f(m)$ for the value on the m th vertex.

Example $f: V \rightarrow R$ s.t. $f_i(m) = i + 1$, $i = 1, \dots\dots, n$

$$\boldsymbol{f} = (\boldsymbol{f}_1, \boldsymbol{f}_2, \boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5) = (2,3,4,5,6)$$

Of key importance for our theory is the graph Laplacian operator $\mathcal{L}$,

Definition 1.2.18.[20]

The non-normalized Laplacian is defined as $\mathcal{L} = D - A$. It can be verified that for any $f \in R^N$, $\mathcal{L}$ satisfies

$$(\mathcal{L}f)(m) = \sum_{m-n} a_{m,n}\left(f(m) - f(n)\right).$$

For example : Graph in Figure 1.7

$$\mathcal{L} = D - A = \begin{pmatrix} 2 & -1 & -1 & 0 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ -1 & -1 & 3 & -1 & 0 \\ 0 & -1 & -1 & 3 & -1 \\ 0 & -1 & 0 & -1 & 2 \end{pmatrix}$$

$$\mathcal{L} - \lambda \mathrm{I} = 0 \qquad , \qquad \mathcal{L}\chi = \lambda\chi$$

Such that, λ: eigenvalues and χ: eigenvectors

Some authors define and use an alternative, normalized form of the Laplacian, defined as,

Definition 1.2.19.[20]

The normalized Laplacian matrix is given by

$$\mathcal{L}^{norm} = \mathrm{D}^{-1/2}\mathcal{L}\,\mathrm{D}^{-1/2} = \mathrm{I} - \mathrm{D}^{-1/2}\,\mathrm{A}\mathrm{D}^{-1/2}$$

It should be noted that $\mathcal{L}$ and $\mathcal{L}^{norm}$ are not similar matrices, in particular their eigenvectors are different. As we shall see in detail later, both operators may be used to define spectral graph wavelet transforms, however the resulting transforms will not be equivalent. Unless noted otherwise we will use the non-normalized form of the Laplacian, however much of the theory presented in this paper is identical for either choice. We consider that the selection of the appropriate Laplacian for a particular problem should depend on the application at hand.

Entries are:

$$\mathcal{L}^{norm} = \begin{cases} 1 & if\ \ i = j \\ -\dfrac{1}{\sqrt{deg(v_i)deg(v_j)}} & if\ i \neq j\ \ and\ v_i\ is\ adjacent\ to\ v_j \\ 0 & otherwise \end{cases}$$

For example : Graph in Figure 1.7

$$\mathcal{L}^{norm} = \begin{pmatrix} 1 & -1/\sqrt{8} & -1/\sqrt{6} & 0 & 0 \\ -1/\sqrt{8} & 1 & -1/\sqrt{12} & -1/\sqrt{12} & -1/\sqrt{8} \\ -1/\sqrt{6} & -1/\sqrt{12} & 1 & -1/3 & 0 \\ 0 & -1/\sqrt{12} & -1/3 & 1 & -1/\sqrt{6} \\ 0 & -1/\sqrt{8} & 0 & -1/\sqrt{6} & 1 \end{pmatrix}$$

Definition 1.2.20.[20]

In order to formulate the desired localization result, we must specify a notion of distance between points m and n on a weighted graph. We will use the shortest-path distance, i.e. the minimum number of edges for any paths connecting m and n:

$$d_G(m,n) = \underset{s}{\operatorname{argmin}}\{k_1, k_2, \dots, k_s\}$$

s.t. $m = k_1\ ,\ n = k_s\ ,\ and\ \ a_{k_r,k_{r+1}} > 0\ \ for\ \ 1 \leq r < s$

For example : Graph in Figure 1.7

$$d_G(v_1,v_2) = 1\ ,\ d_G(v_1,v_3) = 1\ ,\ d_G(v_1,v_4) = 2\ ,\ d_G(v_1,v_5) = 2$$

$$d_G(v_2,v_1) = 1\ ,\ d_G(v_2,v_3) = 1\ ,\ d_G(v_2,v_4) = 1\ ,\ d_G(v_2,v_5) = 1$$

$$d_G(v_3,v_1) = 1\ ,\ d_G(v_3,v_2) = 1\ ,\ d_G(v_3,v_4) = 1\ ,\ d_G(v_3,v_5) = 2$$

$$d_G(v_4,v_1) = 2\ ,\ d_G(v_4,v_2) = 1\ ,\ d_G(v_4,v_3) = 1\ ,\ d_G(v_4,v_5) = 1$$

$d_G(v_5, v_1) = 2$, $d_G(v_5, v_2) = 1$, $d_G(v_5, v_3) = 2$, $d_G(v_5, v_4) = 1$

Note that as we have defined it, d_G disregards the values of the edge weights.

1.3. Wavelets [20]

We first give an overview of the classical continuous wavelet transform (CWT) for $L^2(R)$.

In general, the CWT is generated by the choice of a single "mother" wavelet ψ. Wavelets at different locations and spatial scales are formed by translating and scaling the mother wavelet. We write this by

$$\psi_{s,a}(x) = \frac{1}{s}\psi\left(\frac{x-a}{s}\right)$$

This scaling convention preserves the L^1 norm of the wavelets. Other scaling conventions are common, especially those preserving the L^2 norm, however in our case the L^1 convention will be more convenient.We restrict ourselves to positive scales $s > 0$.

Definition 1.3.1.[20]

For a given signal f , the wavelet coefficient at scale s and location a is given by the inner product of f with the wavelet $\psi_{s,a}$, i.e.

$$w_f\ (s,a)\ = \int_{-\infty}^{\infty} \frac{1}{s}\psi^*\left(\frac{x-a}{s}\right) f(x)dx$$

The CWT may be inverted provided that the wavelet ψ satisfies the admissibility condition

$$\int_0^{\infty} \frac{|\psi(w)|^2}{w} dw = C_\psi < \infty$$

This condition implies, for continuously differentiable ψ, that

$$\hat{\psi}(0)\ = \int \psi(x)dx =\ 0,$$

So ψ must be zero mean. Inversion of the CWT is given by the following relation [25]

$$f(x) = \frac{1}{C_\psi}\int_{0}^{\infty}\int_{-\infty}^{\infty} w_f(s,a)\psi_{s,a}(x)\frac{da\,ds}{s}$$

Translation of the wavelets may be defined through "localizing" the wavelet operator by applying it to an impulse. Writing

$$\delta_a(x) = \delta(x-a),$$

We have,

$$(T^s\delta_a)(x) = \frac{1}{s}\psi^*\left(\frac{a-x}{s}\right)$$

For real-valued and even ψ this reduces to

$$(T^s\delta_a)(x) = \psi_{s,a}(x).$$

In [21], authors defined spectral graph continuous wavelet Transform , in the next chapter, we define L_p spectral graph wavelet transforms and study their properties in details. Note that the graph we use in this research is the simple undirected graph, which implies a symmetric adjacency matrix . moreover . we consider non-negaitive weights for the graph .

Chapter Two

L_p Wavelet Transform

In previous Chapter, particularly in section 1.3, continuous wavelet transform has been introduced in $L^2(R)$[24]. Now, we define $L_p, p < 1$ wavelet transform, for a given function defined on the vertices of a weighted graph, we need to define $\psi(sx)$. For this purpose, we need to go back to fourier transform to construct a basis for the spectral graph wavelet transform .

2.1. Graph Fourier Transform

A 2π periodic function $f(x)$ is the sum

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

of its fourier series. The coefficients a_0, a_k and b_k are calculated by

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)dx, \quad a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos(kx)\,dx, \quad b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin(kx)\,dx\,.[20]$$

As the graph Laplacian $\mathcal{L}$ is a real symmetric matrix, it has a complete set of orthonormal eigenvectors. We denote these by χ_ℓ for $\ell = 0, \dots, N-1$, with associated eigenvalues λ_ℓ

$$\mathcal{L}\chi_\ell = \lambda_\ell \chi_\ell$$

As $\mathcal{L}$ is symmetric, each of the λ_ℓ's are real. For the graph Laplacian, it can be shown that the eigenvalues are all nonnegative, and that 0 appears as an eigenvalue with multiplicity equal to the number of connected components of the graph [25]. Henceforth, we assume the graph G to be connected, we may thus order the eigenvalues such that

$$0 = \lambda_0 < \lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{N-1}$$

For any function $f \in R^N$ defined on the vertices of G, its graph fourier transform $\hat{f}$ is defined by

$$\hat{f}(\ell) = \sum_{n=1}^{N} \chi_\ell^*(n) f(n) \quad , \quad \ell = 0, \dots, N-1$$

Where we adopt the convention that the inner product be conjugate-linear in the first argument. The inverse transform reads as

$$f(n) = \sum_{\ell=0}^{N-1} \hat{f}(\ell) \chi_\ell(n) \, , \quad n = 1, \dots, N$$

2.2. L_p Spectral Graph Wavelet Transform

L_p SGWT can be defined in term of the choice of a kernal function $g: R^+ \to R^+$ that satisfies

$$g(0) = 0 \;\; and \;\; \lim_{x \to \infty} g(x) = 0 \; ,$$

In particular

$$\widehat{T_g f}(\ell) = g(\lambda_\ell) \hat{f}(\ell) \quad \text{where } T_g = g(\ell)$$

With inverse fourier transform ,

$$T_g f(n) = \sum_{\ell=0}^{N} g(\lambda_\ell) \hat{f}(\ell) \chi_\ell(n)$$

Wavelet at scale t is

$$T_g{}^t = g(t\mathcal{L})$$

Applying the above operators at each vertex gives

$$\psi_{t,n} = T_g^t \, \delta_n,$$

Now, L_p SGWT in L_p space, with $p < 1$, is given by

$$\psi_{t,n}(m) = \sum_{\ell=0}^{N} |g(t\lambda_\ell)\chi_\ell^*(n)\chi_\ell\,(m)\,|^p \qquad (2.1)$$

Where

$\psi_{t,n}$: mother wavelet

$n, m = 1, \ldots, N$

λ_ℓ: eigenvalues

χ_ℓ : eigenvectors

g: kernal function

and so we get

$$w_f(t,n) = \left(T_g^t f\right)(n) = \sum_{\ell=0}^{N} \left|g(\lambda_\ell)\hat{f}(\ell)\chi_\ell(n)\,\right|^p \qquad (2.2)$$

Where

$w_f(t,n)$: Wavelet Coefficient with $f \in L_p$

$\hat{f}$: Graph Fourier Transform

$\left(T_g^t f\right)(n)$: Inverse Fourier Transform

2.3. Properties of L_p SGWT

In this section, some properties are studied for L_p SGWT. We first show an inverse formula for the transform under admissible condition.

2.3.1. L_p SGWT Inverse

In order for a particular transform to be useful for signal processing, and not simply signal analysis, it must be possible to reconstruct a signal corresponding to a given set of transform coefficients.

Theorem 2.3.2.

Suppose that L_p SGWT with kernel g satisfies the admissibility condition

$$\int_0^\infty \frac{|g(x)|^{2p}}{x} dx = C_g < \infty$$

and $g(0) = 0$, then

$$\frac{1}{C_g(p)} \sum_{n=1}^{N-1} \int_0^\infty w_f(t,n)\psi_{t,n}(m) \frac{dt}{t} \le f(m) - \hat{f}(0)\chi_0(m)$$

Proof

By (2.1) , (2.2) and admissibility condition, we have

$$\frac{1}{C_g} \sum_{n=1}^{N-1} \int_0^\infty w_f(t,n)\psi_{t,n}(m) \frac{dt}{t}$$

$$= \frac{1}{C_g} \int_0^\infty \frac{1}{t} \sum_n \left(\sum_\ell \left| g(t\lambda_\ell)\hat{f}(\ell)\chi_\ell(n) \right|^p \sum_{\ell'} \left| g(t\lambda_{\ell'})\chi_{\ell'}^*(n)\chi_{\ell'}(m) \right|^p \right) dt$$

$$= \frac{1}{C_g} \int_0^\infty \frac{1}{t} \left(\sum_{\ell,\ell'=0} \left| g(t\lambda_{\ell'})g(t\lambda_\ell)\hat{f}(\ell)\chi_{\ell'}(m) \right|^p \sum_n \left| \chi_{\ell'}^*(n)\chi_\ell(n) \right|^p \right) dt$$

$$\le \frac{C(p)}{C_g} \int_0^\infty \frac{1}{t} \left(\sum_{\ell,\ell'=0} \left| g(t\lambda_{\ell'})g(t\lambda_\ell)\hat{f}(\ell)\chi_{\ell'}(m) \right|^p \sum_n \chi_{\ell'}^*(n)\chi_\ell(n) \right) dt$$

Since of orthonormality of the χ_ℓ, we get

$$\frac{1}{C_g} \sum_{n=1}^{N-1} \int_0^\infty w_f(t,n)\psi_{t,n}(m) \frac{dt}{t}$$

$$\leq \frac{C(p)}{C_g}\left\{\sum_{\ell=1}^{N-1}\left(\int_0^\infty \frac{|g^2(t\lambda_\ell)|^p}{t}dt\right)\hat{f}(\ell)\chi_\ell(m)+\hat{f}(0)\chi_0(m)\right\}$$

Now, substitute $t\lambda_\ell = u$, *to* simplify

$$\int_0^\infty \frac{|g^2(t\lambda_\ell)|^p}{t}dt,$$

Then we get the admissibility condition with $\lambda_\ell dt = du$,

$$\int_0^\infty \frac{|g^2(t\lambda_\ell)|^p}{t}dt = \int_0^\infty \frac{|g(u)|^{2p}}{\frac{u}{\lambda_\ell}}\frac{du}{\lambda_\ell} = \int_0^\infty \frac{|g(u)|^{2p}}{u}du = C_g$$

Since $g(0) = 0$, and

$$\frac{1}{C_g}\sum_{\ell=0}^{N-1}\left(\int_0^\infty \frac{|g(t\lambda_\ell)|^{2p}}{t}dt\right)\hat{f}(\ell)\chi_\ell(m)$$

$$= \frac{1}{C_g}\left\{\sum_{\ell=1}^{N-1}\left(\int_0^\infty \frac{|g(t\lambda_\ell)|^{2p}}{t}dt\right)\hat{f}(\ell)\chi_\ell(m)+\hat{f}(0)\chi_0(m)\right\}$$

$$= f(m) - \hat{f}(0)\chi_0(m) \rightarrow f(m) = \hat{f}(0)\chi_0(m)$$

2.3.3.Localization in small-scale limit

One of the primary motivations for the use of wavelets is that they provide simultaneous localization in both frequency and time (or space). It is clear by construction that if the kernel g is localized in the spectral domain, then the associated spectral graph wavelets will all be localized in frequency. In order to be able to claim that the spectral graph wavelets can yield localization in both frequency and space, however, we must analyze their behavior in the space domain more carefully. For the classical wavelets on the real line, the space localization is readily apparent:

If the mother wavelet $\psi(x)$ is well localized in the interval $[-\varepsilon, \varepsilon\,]$, then the wavelet $\psi_{t,a}(x)$ will be well localized within $[a - \varepsilon\, t, a + \varepsilon t]$. In particular, in the limit as $t \to 0$, $\psi_{t,a}(x) \to 0$ for $x \neq a$. The situation for the spectral graph wavelets is less straightforward to analyze because the scaling is defined implicitly in the Fourier domain. We will nonetheless show that, for g sufficiently regular near 0, the normalized spectral graph wavelet $\frac{\psi_{t,j}}{\|\psi_{t,j}\|_p}$ will vanish on vertices sufficiently far from j in the limit of fine scales, i.e. as t → 0. This result will provide a quantitative statement of the localization properties of the spectral graph wavelets. One simple notion of localization for $\psi_{t,n}$ is given by its value on a distant vertex m, e.g. we should expect $\psi_{t,n}(m)$ to be small if n and m are separated, and t is small. Note that

$$\psi_{t,n}(m) = \sum \left|\psi_{t,n}\delta_m\right|^p = \sum \left|T_g^t \delta_n \delta_m\right|^p$$

In order to transfer the study of the localization property from g to an approximating polynomial, we need to examine the stability of the wavelets under perturbations of the generating kernel. This, together with the Taylor approximation will allow us to examine the localization properties for integer powers of the Laplacian $\mathcal{L}$.

Lemma 2.3.4.[20]

Let G be a weighted graph, $\mathcal{L}$ the graph Laplacian (normalized or non-normalized) and $s > 0$ an integer. For any two vertices m and n, if $d_G(m, n) > s$ then $(\mathcal{L}^s)_{m,n} = 0$.

We now proceed to examining how perturbations in the kernel g affect the wavelets in the vertex domain. If two kernels g and $\tilde{g}$ are close to each other in some sense, then the resulting wavelets should be close to each other. More precisely, we have

Theorem 2.3.5.

Let $\psi_{t,n} = T_g^t \delta_n$ and $\tilde{\psi}_{t,n} = T_{\tilde{g}}^t \delta_n$ be the wavelets at scale t generated by the kernels g and $\tilde{g}$. If $\|g(t\lambda) - \tilde{g}(t\lambda)\|_p \le C(t)$ for all $\lambda \in [0, \lambda_{N-1}]$, then $\left\|\psi_{t,n}(m) - \tilde{\psi}_{t,n}(m)\right\|_p \le C(t, p)$ for each vertex m.

Proof

First recall (2.1)

$$\psi_{t,n}(m) = \sum_{\ell} |\chi_\ell(m) g(t\lambda_\ell) \chi_\ell^*(n)|^p$$

Thus,

$$\left\|\psi_{t,n}(m) - \tilde{\psi}_{t,n}(m)\right\|_p^p$$

$$= \left\|\sum_{\ell} |\chi_\ell(m) g(t\lambda_\ell) \chi_\ell^*(n)|^p - \sum_{\ell} |\chi_\ell(m) \tilde{g}(t\lambda_\ell) \chi_\ell^*(n)|^p\right\|_p^p$$

$$= \sum_{\ell} \left|\left|\sum_{\ell} \chi_\ell(m)(g(t\lambda_\ell) - \tilde{g}(t\lambda)) \chi_\ell^*(n)\right|^p\right|^p$$

$$\le C(t, p) \sum_{\ell} |\chi_\ell(m) \chi_\ell(n)^*|^{p^2}$$

$$\leq C(t,p) \sum_{\ell} |\chi_\ell (m)\chi_\ell(n)^*|^p$$

$$\leq C(t,p)$$

We prove the final localization result for kernels g which have a zero of integer multiplicity at the origin. Such kernels can be approximated by a single monomial for small scales..

Theorem 2.3.6.

Let $g \in L_p^{K+1}$, g be $K+1$ times continuously differentiable, satisfying $g(0) = 0$, $g^{(r)}(0) = 0 \, for \, all \;\; r < K \, , and \; g^{(K)}(0) = C \neq 0.$ Assume that there is some $t_1 > 0$ such that $\|g(t\lambda)\|_p \leq B$ for all $\lambda \in [0, t_1\lambda_{N-1}]$. Then, for $\tilde{g}(t\lambda) = (C/K!)(t\lambda)^K$ we have

$$C(t) = \|g(t\lambda) - \tilde{g}(t\lambda)\|_p \leq B\frac{t^{K+1}\lambda_{N-1}^{K+1}}{(K+1)!}$$

for all $t < t_1$.

Proof.

As the first $K-1$ derivatives of g are zero, Taylor's formula with remainder shows, for any values of t and λ,

$$g(t\lambda) = C\frac{(t\lambda)^K}{K!} + g^{(K+1)}(x^*)\frac{(t\lambda)^{K+1}}{(K+1)!}$$

For some $x^* \in [0, \mathrm{t}\lambda]$. Now fix $t < t_1$. For any $\lambda \in [0, \lambda_{N-1}]$, we have $t\lambda < t_1\lambda_{N-1}$, and so the corresponding $x^* \in [0, t_1\lambda_{N-1}]$, and so $\|g(t\lambda)\|_p \leq B$. This implies

$$\|g(t\lambda) - \tilde{g}(t\lambda)\|_p \leq \left\| C\frac{(t\lambda)^K}{K!} + g^{(K+1)}(x^*)\frac{(t\lambda)^{K+1}}{(K+1)!} - \frac{C}{K!}(t\lambda)^K(t\lambda) \right\|_p$$

$$B\frac{t^{K+1}\lambda^{K+1}}{(K+1)!} \leq B\frac{t^{K+1}\lambda_{N-1}^{K+1}}{(K+1)!}$$

As this holds for all $\lambda \in [0, \lambda_{N-1}]$

We are now ready to state the complete localization result. Note that due to the normalization chosen for the wavelets, in general $\psi_{t,n}(m) \to 0$ as $t \to 0$ for all m and n. Thus a non-vacuous statement of localization must include a renormalization factor in the limit of small scales.

Theorem 2.3.7.

Let G be a weighted graph with Laplacian $\mathcal{L}$. Let g be a kernel satisfying the hypothesis of Theorem(2.3.6), with constants t_1 and B. Let m and n be vertices of G such that $d_G(m,n) > K$. Then there exist constants D and t, such that

$$\frac{\psi_{t,n}(m)}{\|\psi_{t,n}\|_p} \leq Dt$$

for all $t < \min(t_1, t_2)$.

Proof

Set $\tilde{g}(\lambda) = \frac{g^{(K)}(0)}{K!}\lambda^K$ and $\tilde{\psi}_{t,n} = T_{\tilde{g}}^{t}\delta_n$. We have

$$\tilde{\psi}_{t,n}(m) = \frac{g^{(K)}(0)}{K!}t^K \sum |\tilde{g}(t\lambda_\ell)\chi_\ell^*(n)\chi_\ell(m)|^p = 0$$

By Lemma (2.3.4) , as $d_G(m,n) > K$. By the results of Theorems (2.3.5).and (2.3.6).we have

$$\psi_{t,n}(m) - \tilde{\psi}_{t,n}(m) = \psi_{t,n}(m) \leq t^{K+1}\frac{\lambda_{N-1}^{K+1}}{(K+1)!}B \qquad (2.3)$$

Writing

$$\psi_{t,n} = \tilde{\psi}_{t,n} + (\psi_{t,n} - \tilde{\psi}_{t,n})$$

and applying the quasi-triangle inequality shows

$$\left\|\tilde{\psi}_{t,n}\right\|_p - \left\|\psi_{t,n} - \tilde{\psi}_{t,n}\right\|_p \le C\left\|\psi_{t,n}\right\|_p \qquad (2.4)$$

We may directly calculate

$$\left\|\tilde{\psi}_{t,n}\right\|_p = t^K \frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p$$

and we have from Theorem **(2.3.6)**

$$\left\|\psi_{t,n} - \tilde{\psi}_{t,n}\right\|_p \le t^{K+1} \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B$$

These imply together that the l.h.s. of (2.4) is greater than or equal to

$$\left\|\psi_{t,n} - \tilde{\psi}_{t,n}\right\|_p \le t^K \left(\frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p - t \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B \right)$$

But (2.4) gives

$$\left\|\psi_{t,n}\right\|_p \ge C\left\|\psi_{t,n} - \tilde{\psi}_{t,n}\right\|_p$$

Together with (2.3), this shows

$$\frac{\psi_{t,n}(m)}{\left\|\psi_{t,n}\right\|_p} \le C \frac{t \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B}{\frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p - t \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B}$$

An elementary calculation shows

$$\frac{\frac{\lambda_{N-1}^{K+1}}{(K+1)!} t}{\frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p - t \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B} \le \frac{2 \frac{\lambda_{N-1}^{K+1}}{(K+1)!}}{\frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p} t$$

If

$$t \le C \frac{\frac{g^{(K)}(0)}{K!} \|\mathcal{L}^K \delta_n\|_p}{2 \frac{\lambda_{N-1}^{K+1}}{(K+1)!} B}$$

This implies the desired result with

$$D = C \frac{2 \frac{\lambda_{N-1}^{K+1}}{(K+1)!} BK!}{g^{(K)}(0) \, \|\mathcal{L}^K \delta_n\|_p} \quad \text{and} \quad t_2 = \frac{g^{(K)}(0) \, \|\mathcal{L}^K \delta_n\|_p (K+1)}{2\lambda_{N-1}^{K+1} B}$$

Theorem 2.3.8.

Given a set of scales $\{t_j\}_{j=1}^{J}$, the set $F = \{\varphi_n\}_{n=1}^{N} \cup \{\psi_{t\ j\ ,n}\}_{j=1}^{J}{}_{n=1}^{N}$ forms a frame

$$A\|f\|_p \leq \left\|W_f(t,n)\right\|_p \left\|S_f(n)\right\|_p \leq B\|f\|_p$$

For any scaling function coefficients $S_f(n)$, with kernel h and bounds A, B given by

$$A = \min_{\lambda \in [0, \lambda_{N-1}]} G(\lambda) \text{, and}$$

$$B = \max_{\lambda \in [0, \lambda_{N-1}]} G(\lambda) \text{,}$$

Where

$$G(\lambda) = |h(\lambda_\ell)|^p + \sum_j \left|g(t_j\, \lambda)\right|^p$$

Proof

Fix f . Using expression (2.2), we see

$$\left\|w_f(t,n)\right\|_p^p = \sum_{n=1}^{N} \left|w_f(t,n)\right|^p$$

$$= \sum_{n=1}^{N} \sum_{\ell=0}^{N-1} \left|g(t\lambda_\ell)\chi_\ell(n)\hat{f}(\ell)\right|^p$$

$$\leq C \sum_{\ell=0}^{N-1} |g(t\lambda_\ell)|^p \left|\hat{f}(\ell)\right|^p \qquad (2.5)$$

Similarly, for any scaling function coefficients $S_f(n)$, we have

$$\left\|S_f(n)\right\|_p^p = \sum_{n=1}^{N} \left|S_f(t,n)\right|^p$$

$$= \sum_{n=1}^{N} \left|h(\lambda_\ell)\chi_\ell(n)\hat{f}(\ell)\right|^p$$

$$\leq C \sum_{\ell=0}^{N-1} |h(\lambda_\ell)|^p \left|\hat{f}(\ell)\right|^p \qquad (2.6)$$

Using (2.5) and (2.6), we have

$$Q = C \sum_{\ell=0}^{N-1} \left(|h(\lambda_\ell)|^p + \sum_{j=1}^{J} \left|g\left(t_j\lambda_\ell\right)\right|^p \right) \left|\hat{f}(\ell)\right|^p = C \sum_{\ell=0}^{N-1} G(\lambda_\ell) \left|\hat{f}(\lambda_\ell)\right|^p$$

Then by the definition of A and B, we have

$$A \sum_{\ell=0}^{N-1} \left|\hat{f}(\ell)\right|^p \leq Q \leq B \sum_{\ell=0}^{N-1} \left|\hat{f}(\ell)\right|^p$$

Chapter Three

Approximatrion by L_pSGWT, Application in Approximation Theory

In this chapter , we benefit from the construction of L_pSGWT to find a best approximation of vertex functions out of space of L_pSGWT. As the approximation began with polynomials, we approximate the kernel g with a polynomial p in a way that leads to the approximation of their corresponding wavelet operators in L_p spaces. The following theorem shows that the degree of approximation between the two wavelets is at most the degree of approximation between their generators.

3.1. Polynomial Approximation

Theorem 3.1.1.

Let $\lambda_{max} \geq \lambda_{N-1}$ be any upper bound on the spectrum of $\mathcal{L}$. For fixed $t > 0$, let $P(x)$ be a polynomial best approximant of $g(tx)$ with degree of approximation $E_n(g)_p = B$ for the space of polynomials of degree at most n. Then the approximate wavelet coefficients $\widetilde{w}_f(t,n) = (P(\mathcal{L})\, f\,)_n$ satisfies

$$\left\|w_f(t,n) - \widetilde{w}_f(t,n)\right\|_p \leq C\|f\|_p$$

Proof

By using (2.2) we get

$$\left\|w_f(t,n) - \widetilde{w}_f(t,n)\right\|_p^p$$

$$= \left\|\sum_\ell \left|g(t\lambda_\ell)\hat{f}(\ell)\chi_\ell(n)\right|^p - \sum_\ell \left|P(t\lambda_\ell)\hat{f}(\ell)\chi_\ell(n)\right|^p\right\|_p^p$$

$$\leq \left\|\sum_\ell \left|\big(g(t\lambda_\ell) - P(\lambda_\ell)\big)\hat{f}(\ell)\chi_\ell(n)\right|^p\right\|_p^p$$

$$\leq C(p)\sum \left\| \left| \left(g(t\lambda_\ell) - P(\lambda_\ell)\right)\hat{f}(\ell)\chi_\ell(n)\right|^p \right\|_p^p$$

$$\leq C(p)\mathrm{B}^p \, \|f\|_p^p$$

The last step follows from the orthonormality of the χ_ℓ

3.2.Existence of Best Approximation

In order to approximate vertex functions by wavelets of spectral graph type, here is a the theorem that confirm the possibility of existence of best approximation of vertex functions out of the space of L_pSGWT, namely Ω.

Theorem 3.2.1.

For any $f \in L_p(R^N)$, the space of N-dimensional Lebesgus, then there exist $w_f \in \Omega$ of the from (2.2) that is generated by a graph G and a kernel g s.t

$$\left\|f - w_f(t,n)\right\|_p < \varepsilon$$

Proof

Set $\widetilde{w}_f \in \Omega$, with a polynomial P, that satisfies Theorem 3.1.1., so that

$$\left\|w_f - \widetilde{w}_f\right\|_p < C(p)\|f\|_p$$

Also, set

$$\left\|f - \widetilde{w}_f(t,n)\right\|_p < \frac{\varepsilon}{2}$$

Which is true by 1.1.9., then by Quasi-Triangle Inequality, we get the desired result

$$\left\|f - w_f(t,n)\right\|_p \leq C\left[\left\|f - \widetilde{w}_f(t,n)\right\|_p + \left\|w_f - \widetilde{w}_f\right\|_p\right]$$

$$\leq C\left[\frac{\varepsilon}{2} + B\|f\|_p\right] < \varepsilon$$

By choosing the costant B, that satisfies

$$B \leq \frac{\varepsilon}{2\|f\|_p} < \varepsilon$$

CONCLUSION

This research studied the best approximation of any function defined on the vertices of a graph with wavelets of spectral graph type. The construction of such transforms was considered in terms of Fourier graph in the space $\boldsymbol{L_p}$. Properties of $\boldsymbol{L_p}$SGWT were studied, such as inversion, localization and scaling.

FUTURE WORK

We defined L_p Spectral Graph Wavelet Transform and got the best approximation in terms of ε . As for future research, we will focus on improving the degree of approximation based on the continuity modulus, to get the result

$$\left\|f - w_f(t,n)\right\|_p \leq C\omega(f,\delta)$$

Also, we can study and prove inverse theorem, in terms of $\omega(f,\delta)$, to get

$$\omega(f,\delta) \leq C\left\|f - w_f(t,n)\right\|_p$$

REFERENCES

[1] F.K. Chung, Spectral Graph Theory, CBMS Regional Conference Series in Mathematics. Number 92, AMS Bookstore, 1997.

[2] S. Mallat, A Wavelet Tour of Signal Processing, Academic Press, 1998.

[3] J. Antoine and P. Vandergheynst, Wavelets on the 2-sphere: A group-theoretical approach, Applied and Computational Harmonic Analysis. 7 .(3) (1999) 262–291.

[4] Y. Wiaux, J.D. McEwen, P. Vandergheynst, O. Blanc, Exact Reconstruction with Directional Wavelets on the Sphere, Monthly Notices of the Royal Astronomical Society, 388 (2008) 770-788

[5] G. Peyré and S. Mallat, Orthogonal Bandlet Bases for Geometric Images Approximation, Communications on Pure and Applied Mathematics. 61 (9) (2008) 1173–1212.

[6] J.-P. Antoine, I. Bogdanova and P. Vandergheynst, The Continuous Wavelet Transform on Conic Sections, International Journal of Wavelets, Multiresolut and Information Processing. 6 (2) (2008) 137–156.

[7] M. Crovella and E. Kolaczyk, Graph Wavelets for Spatial Traffic Analysis, in: Infocom 2003, Twenty-Second Annual

Joint Conference of the IEEE Computerand Communications Societies. 3, 2003. 1848–1857.

[8] A. Smalter, J. Huan, G. Lushington, Graph Wavelet Alignment kernels for drug virtual screening, Journal of Bioinformatics and Computational Biology. 7 (2009) 473–497.

[9] M. Jansen, G.P. Nason, B.W. Silverman, Multiscale Methods for data on Graphs and Irregular Multidimensional Situations, Journal of the Royal Statistical Society: Series B Statistical Methodology. 71 (1) (2009) 97–125.

[10] F. Murtagh, The Haar Wavelet Transform of a Dendrogram, Journal of Classification 24 (1) (2007) 3–32.

[11] A.B. Lee, B. Nadler and L. Wasserman, Treelets an Adaptive Multi-Scale Basis for Sparse Unordered Data, The Annals of Applied. Statistics. 2 (2008) 435–471.

[12] N. L. Carothers, “A Short Course on Approximation Theory,” Math 682 Summer ,Citeseer , 1998.

[13] K. Conrad, “Lp-Spaces for $0 < p < 1$,”, pp. 1–15, 2003.

[14] L.maligranda., “Type, Cotype and Convexity Properties of Quasi–Banach Spaces,”. Proceedings of the International Symposium on Banach and Function Spaces. pp. 83-120, October 2-4, 2003.

[15] F. Deutsch, "Best Approximation in Inner Product Spaces," Canadian Mathematical society (societe Mathematique du Canada) . Springer-Verlag New York, Inc. 2001.

[16] Walter Rudin, "Principles of Mathematical Analysis," Third Edition , Exclusive rights by Mcgraw-Hill Book Co. - Singapore for manufacture and export . International Editions 1976.

[17] T.Roughgarden , G.Valiant, "Spectral Graph Theory," The Modern Algorithmic Toolbox , May 2, 2022

[18] Robin J. Wilson, "Introduction to Graph Theory," . Fourth edition . Produced through Longman Malaysia, PP. Longman Group Ltd, 1998.

[19] X.Chen, "Understanding Spectral Graph Neural Network," Department of Mathematics, University of Manchester Manchester, M13 9PL, United Kingdom , pp. 1–19, 2020.

[20] D.K. Hammond , P.Vandergheynst, R.Gribonval "Applied and Computational Harmonic Analysis,Wavelets on graphs via spectral graph theory" Contents lists available at ScienceDirect, vol. 30, no. 1, pp. 129–150, 2011.

[21] Amara Graps, "An Introduction to Wavelets," Article in IEEE Computational Science and Engineering · February 1995

[22] E. S. Bhaya, "On the Constrained and Unconstrained Approximation", PD Thesis, University of Baghdad, 2003..

[23] A. Grossmann, J. Morlet, Decomposition of Hardy functions into square integrable wavelets of constant shape, SIAM J. Math. Anal. 15 (4) (1984). 723–736.

[24] J. Portilla, V. Strela, M.J. Wainwright, E.P. Simoncelli, Image denoising using scale mixtures of Gaussians in the wavelet domain, IEEE Trans. Image Process. 12 (2003) 1338–1351.

[25] J.-L. Starck, A. Bijaoui, Filtering and deconvolution by the wavelet transform, Signal Process. 35 (3) (1994) 195–211.

المستخلص

يعد تقريب الدوال باستخدام مويجات البيان الطيفي اتجاهًا مثيرًا للاهتمام في نظرية التقريب. ويفتح بابًا لعالم جديد من التقريب الدالي. ومن الضروري اختيار فضاء الدوال التي يتم تقريبها بواسطة مويجات البيان الطيفي بشكل جيد. هذا ويعد فضاء الدوال الليبيكية L_p من الاختيارات الرائعة للدراسة. من المثير للاهتمام أن تأخذ القيمة $0 < p < 1$. في هذا البحث تم انشاء صيغ جديدة لمويجات البيان الطيفي ودراسة خصائصها كالانعكاس والتحجيم وتقريب الموجات لبعضها البعض وفي نهاية المطاف , تم اثبات وجود افضل تقريب للدوال الليبيكية بواسطة موجات البيان الطيفي .

جمهورية العراق

وزارة التعليم العالي والبحث والعلمي

جامعة بابل

كلية التربية للعلوم الصرفة

قسم الرياضيات

تقريب الدوال بواسطة

مويجات البيان الطيفي

بحث مقدم إلى

مجلس كلية التربية للعلوم الصرفة – جامعة بابل

كجزء من متطلبات نيل شهادة الدبلوم العالي تربية / الرياضيات

من قبل

عـــلي سعيــد مناحــي عاجل المشلب

بإشراف

أ.م.د. حـــــوراء عبـــاس فاضـل المرعب

١٤٤٤