An- Najah National University
Faculty of Graduate Studies

Mathematical Theory of Wavelets

By
Bothina Mohammad Hussein Gannam

Supervisor

Dr. Anwar Saleh

Submitted in Partial Fulfillment of the Requirements for the Degree of Master in Science in Mathenatics, Faculty of Graduate Studies, at An-Najah National University, Nablus, Palestine.

2009

Mathematical Theory of Wavelets

By
Bothina Mohammad Hussein

This Thesis was defended successfully on 23/4/2009 and approved by:

Committee Members		**Signature**
1. **Dr. Anwar Saleh**	**Supervisor**	**…………………**
2- **Dr. Samir Matar**	**Internal Examiner**	**………………...…**
3. **Dr. Saed Mallak**	**External Examiner**	**……………………**

Dedication

Dedication to my father and mother
And
To my husband Jihad, and my sons, Abdullah, Muhammad.

Acknowledgement

All praise be to almighty Allah, without whose mercy and clemency nothing would have been possible. I wish to express my appreciation to Dr. Anwar Saleh, my advisor, for introducing me to the subject and also for giving me all the necessary support I needed to complete this work, without him this work would not have been accomplished.

Also, I would like to thank Dr. Saed Mallak , Dr. Samir Matter for their encouragement , support and valuable advice to complete this study.

Acknowledgement is due to An- Najah National University for supporting this research work and in particular, to the Department of Mathematical Science for giving me access to all its available facilities which makes the completion of this work much easier.

إقـــرار

أنا الموقع أدناه مقدم الرسالة التي تحمل العنوان:

Mathematical Theory of Wavelets

اقر بأن ما اشتملت عليه هذه الرسالة إنما هي نتاج جهدي الخاص، باستثناء ما تمت الإشـــارة إليه حيثما ورد، وأن هذه الرسالة ككل، أو أي جزء منها لم يقدم من قبل لنيل أية درجة علمية أو بحث علمي أو بحثي لدى أية مؤسسة تعليمية أو بحثية أخرى.

Declaration

The work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification.

Student's name:	**اسم الطالب:**
Signature:	**التوقيع:**
Date:	**التاريخ:**

Table of contents

Mathematical Theory of Wavelets

By
Bothina Mohammad Hussein Gannam
Supervisor

Dr. Anwar Saleh

Abstract

Wavelets are functions that satisfy certain requirements and are used in representing and processing functions and signals, as well as, in compression of data and images as in fields such as: mathematics, physics, computer science, engineering, and medicine. The study of wavelet transforms had been motivated by the need to overcome some weak points in representing functions and signals by the classical Fourier transforms such as the speed of convergence and Gibbs phenomenon. In addition, wavelet transforms have showed superiority over the classical Fourier transforms. In many applications, wavelet transforms converge faster than Fourier transforms, leading to more efficient processing of signals and data. In this thesis, we overview the theory of wavelet transforms, as well as, the theory of Fourier transforms and we make a comparative theoretical study between the two major transforms proving the superiority of wavelet transforms over the Fourier transforms in terms of accuracy and the speed of convergence in many applications.

Chapter one

Introduction

1.1. A Brief History of Wavelets

1.2. Wavelet

1.3. Applications

1.4. Signal analysis

1.5. Why wavelet?

Chapter 1
Introduction

Wavelets were introduced relatively recently, in the beginning of the 1980s. They attracted considerable interest from the mathematical community and from members of many divers disciplines in which wavelets had promising applications. A consequence of this interest is the appearance of several books on this subject and a large volume of research articles.

The goal of most modern wavelet research is to create asset of basis functions and transforms that will give an informative, efficient, and useful description of a function or signal. If the signal is represented as a function of time, wavelets provide efficient localization in both time and frequency or scale. Another central idea is that multiresolution were the decomposition of a signal is in terms of the resolution of detail.

1.1 A Brief History of Wavelets

In the history of mathematics, wavelet analysis shows many different origins. Much of the work was performed in the 1930s, and, the separate efforts did not appear to be parts of a coherent theory. Wavelets are currently being used in fields such as signal and image processing, human and computer vision, data compression, and many others. Even though the average person probably knows very little about the concept of wavelets, the impact that they have in today's technological world is phenomenal.

The first known connection to modern wavelets dates back to a man named **Jean Baptiste Joseph Fourier**. In 1807, Fourier's efforts with frequency analysis lead to what we know as Fourier analysis. His work is based on the fact that functions can be represented as the sum of sines and cosines.

Another contribution of Joseph Fourier's was the Fourier Transform. It transforms a function *f* that depends on time into a new function which depends on frequency. The notation for the Fourier Transform is indicated below.

$$\hat{f}(w) = \int_{\Re} f(x) e^{-iwx} dx \, .$$

The next known link to wavelets came 1909 from **Alfred Haar** . It appeared in the appendix of a thesis he had written to obtain his doctoral degree. Haar's contribution to wavelets is very evident. There is an entire wavelet family named after him. The Haar wavelets are the simplest of the wavelet family and are easy to understand.

After Haar's contribution to wavelets there was once again a gap of time in research about the functions until a man named **Paul Levy**. Levy's efforts in the field of wavelets dealt with his research with Brownian motion. He discovered that the scale-varying basis function – created by Haar (i.e. Haar wavelets) were a better basis than the Fourier basis functions. Unlike the Haar basis function, which can be chopped up into different intervals – such as the interval from 0 to 1 or the interval from 0 to ½ and ½ to 1, the Fourier basis functions have only one interval. Therefore, the Haar wavelets can be much more precise in modeling a function.

Even though some individuals made slight advances in the field of wavelets from the 1930's to 1970's, the next major advancements came from **Jean Morlet** around the year 1975. In fact, Morlet was the first researcher to use the term "wavelet" to describe his functions. More specifically, they were called "Wavelets of Constant Slope".

Morlet had made quite an impact on the history of wavelets; however, he wasn't satisfied with his efforts by any means. In 1981, Morlet teamed up with a man named **Alex Grossman**. Morlet and Grossman worked on the idea that Morlet discovered while experimenting on a basic calculator. The idea was that a signal could be transformed into wavelet form and then transformed back into the original signal without any information being lost. When no information is lost in transferring a signal into wavelets and then back, the process called lossless. Since wavelet deal with both time and frequency, they thought a double integral would be needed to transform wavelet coefficients back into the original signal. However, in 1984, Grossman found that a single integral was all that was needed.

While working on this idea, they also discovered another interesting thing. Making a small change in the wavelets only causes a small change in the original signal. This is also used often with modern wavelets. In data compression, wavelet coefficients are changed to zero to allow for more compression and when the signal is recomposed the new signal is only slightly different from the original.

The next two important contributors to the field of wavelets were **Yves Meyer** and **Stephane Mallat**. In 1986, Meyer and Mallat first formulated the idea of multiresolution analysis (MRA) in the context of wavelet analysis. This idea of multiresolution analysis was a big step in the research of wavelets, which deals with a general formalism for construction of an orthogonal basis of wavelets. Indeed, (MRA) is a central to all constructions of wavelet bases.

A couple of years later, **Ingrid Daubechies**, who is currently a professor at Princeton University, used Mallat's work to construct a set of wavelet orthonormal basis functions, and have become the cornerstone of wavelet applications today.

1.2 Wavelet

A wave is usually defined as an oscillation function of time or space, such as a sinusoid. Fourier analysis is wave analysis. It expands signals or functions in terms of sines and cosines which has proven to be extremely valuable in mathematics, science, and engineering, especially for periodic, time-invariant, or stationary phenomena. A wavelet is a "small wave", which has its energy concentrated in time to give a tool for the analysis of transient, nonstationary phenomena.

A reason for the popularity of wavelet is its effectiveness in representation of nonstationary (transient) signals. Since most of natural and human-made signals are transient in nature, different wavelets have been used to represent this much larger class of signals than Fourier representation of stationary signals. Unlike Fourier- based analyses that use global (nonlocal) sine and cosine functions as bases, wavelet analysis uses bases that are localized in time and frequency to represent nonstationary signals more effectively. As a result, a wavelet representation is much more compact and easier to implement. Using the powerful multiresolution analysis, one can represent a signal by a finite sum of components at different resolutions so that each component can be processed adaptively based on the objectives of the application. This capability to represent signals compactly and in several levels of resolution is the major strength of wavelet analysis.

1.3 Applications

Wavelet analysis is an exiting new method for solving difficult problems in mathematics, physics, and engineering, with modern applications as diverse as wave propagation, data compression, image processing, pattern recognition, computer graphics, the detection of aircraft and submarines, and improvement in CAT scans and other medical image technology. Wavelets allow complex information such as music, speech, images, and patterns to be decomposed in to elementary forms, called the *fundamental building blocks*, at different positions and scales and subsequently reconstructed with high precision.

1.4 Signal analysis

Fourier analysis and the wavelet analysis play the major role in signal processing. In fact, large part of the development of such transforms is due to their role in signal processing. In this section, we give a short overview of signals. Signals are categorized in two ways: Analog signals and Discrete signals.

Definition 1.3.1 [8]: Analog Signals

An analog signal is a function $X:\Re \to \Re$, where $\Re$ is the set of real numbers, and $X(t)$ is the signal value at time *t*.

Example 1.3.1: Unit step signal

The unit step signal $X(t)$ is defined by:

$$X(t) = \begin{cases} 1 & \text{if } t \geq 0 \\ 0 & \text{if } t < 0 \end{cases}$$

and it is a building block for signals that consist of rectangular shapes and square pulses.

Unlike analog signals, which have a continuous domain, the set of real numbers $\Re$, discrete signals take values on the set of integers Z. Each integer *n* in the domain of *x* represents a time instant at which the signal has a value *x* (*n*).

Definition 1.3.2 [8]: Discrete and Digital Signals

A *discrete-time* signal is a real-valued function $x:\mathrm{Z} \to \Re$, with domain is the set of integer set Z. $x(n)$ is the signal value at time instant *n*. A *digital signal* is an integer-valued function $x:\mathrm{Z} \to [-N, N]$, with domain Z, and $N \in \mathrm{Z}$, $N > 0$.

Example 1.3.2: Discrete Unit step

The unit step signal $x(n)$ is defined by:

$$x(n) = \begin{cases} 1 & \text{if } n \geq 0 \\ 0 & \text{if } n < 0 \end{cases}$$

The most important signal classes are the discrete and analog finite energy signals.

Definition 1.3.3 [8]: Finite-Energy Discrete Signals
A discrete signal $x(n)$ has finite-energy if $\sum_{n\in Z}|x(n)|^2 < \infty$

Definition 1.3.4 [8]: Finite-Energy Analog Signals
An analog signal $X(t)$ is finite-energy if $\int_{\Re}|X(t)|^2 < \infty$

The term" finite-energy" has a physical meaning. The amount of energy required to generate a real-world signal is proportional to the total squares of its values.

1.5 Why wavelet?

One disadvantage of Fourier series is that its building blocks, sines and cosines, are periodic waves that continue forever. While this approach may be appropriate for filtering or compressing signals that have time-independent wavelike features, other signals may have more localized features for which sines and cosines do not model very well. A different set of building blocks, called wavelets, is designed to model these types of signals.

Another shortcoming of Fourier series exists in convergence. In 1873, **Paul Du Bois-Reymond** constructed a continuous, 2π-periodic function, whose Fourier series diverge at a given point. Many years later Kolmogorove (1926) had proved the existence of an example of 2π-periodic, L^1 function has Fourier series diverged at every point. This raised the question of convergence of Fourier series and motivated mathematicians to think of other possible orthogonal system that is suitable for any 2π-periodic function by avoiding divergence of the Fourier series representation.

This thesis consists of three chapters. In chapter 2, the basics of Fourier series and several convergence theorems are presented with simplifying hypothesis so that their proofs are manageable. The Fourier transform is also presented with a formal proof of the Fourier inversion formula. Several important results including the convolution theorem, parseval's relation, and various summability kernels are discussed in some detail. Included are Poisson's summation formula, Gibbs's phenomenon, the Shannon sampling theorem.

Chapter 3 is devoted to wavelets and wavelet transforms with examples. The basic ideas and properties of wavelet transforms are mentioned. In addition, the formal proofs for the parseval's and the inversion formulas for the wavelet transforms are presented. Our presentation of wavelets starts with the case of the Haar wavelets. The basic ideas behind a multiresolution analysis and desired features of wavelets, such as orthogonality, are easy to describe with the explicitly defined Haar wavelets. Finally, some convergence theorems for the wavelet series are presented.

In chapter 4, the speed of convergence for Fourier and wavelet series by studying the rate of decay for those coefficients have been discussed. At the end of this chapter we set some differences between the Fourier and wavelet transforms.

Chapter two

Fourier Analysis

2.1. Introduction

2.2. Fourier series

2.3. Functional spaces

2.4. Convergence of Fourier series

2.5. Summability of Fourier series

2.6. Generalized Fourier series

2.7. Fourier Transform

Chapter 2
Fourier Analysis

2.1 Introduction

Historically, **Joseph Fourier** (1770-1830) first introduced the remarkable idea of expansion of a function in terms of trigonometric series without rigorous mathematical analysis. The integral formulas for the coefficients were already known to **Leonardo Euler** (1707-1783) and others. In fact, Fourier developed his new idea for finding the solution of heat equation in terms of Fourier series so that the Fourier series can be used as a practical tool for determining the Fourier series solution of a partial differential equation under prescribed boundary conditions.

The subject of Fourier analysis (Fourier series and Fourier transform) is an old subject in mathematical analysis and is of great importance to mathematicians, scientist, and engineers alike. The basic goal of Fourier series is to take a signal, which will be considered as a function of time variable t, and decompose into various frequency components. In other words, transform the signal from time domain to frequency domain, so it can be analyzed and processed. As an application is the digital signal processing. The basic building blocks are the sine and cosine functions, which vibrate at frequency of n times per 2π intervals.

2.2 Fourier series

Fourier series is a mathematical tool used to analyze periodic functions by decomposing such functions into sum of simple functions, which may be sines and cosines or may be exponentials.

Definition 2.2.1 [24]: Fourier series

If f is periodic function with period 2π and is integrable on $[-\pi,\pi]$, then the Fourier series expansion of f is defined as: $a_0+\sum_{n=1}^{\infty} a_n\cos nx+b_n\sin nx$, where the coefficients $a_0, a_n, b_n, (n\in \mathbf{Z}^+)$ in this series, called the Fourier coefficients of f, are defined by:

$$a_0=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)dx \tag{2.2.1}$$

$$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos nxdx \tag{2.2.2}$$

$$b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin nxdx \tag{2.2.3}$$

This definition can be generalized to include periodic functions with period $p=2L$, for any positive real number L, by using the trigonometric functions $\cos\left(\frac{n\pi x}{L}\right)$, $\sin\left(\frac{n\pi x}{L}\right)$ and the following lemma.

Lemma 2.2.2 [4]: Suppose f is any 2π-periodic function and c is any real number, Then

$$\int_{-\pi+c}^{\pi+c} f(x)dx=\int_{-\pi}^{\pi} f(x)dx$$

The following theorem illustrates the generalization of Fourier series to functions of any period.

Theorem 2.2.3 [4]: If $f(x)=a_0+\sum_{n=1}^{\infty} a_n\cos\left(\frac{n\pi x}{L}\right)+b_n\sin\left(\frac{n\pi x}{L}\right)$ on the interval $[-L,L]$, then

$$a_0 = \frac{1}{2L}\int_{-L}^{L} f(x)dx$$
$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx$$
$$b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right)dx$$

One major application of Fourier series is in signal analysis where signals are analyzed and processed. Many signals are periodic or symmetric. In fact, any signal can be decomposed into an even part and odd part, where analysis can be easier.

Theorem 2.2.4 [4]: Suppose f is a periodic function with period $p = 2L$ defined on the interval $[-L, L]$.

a. If f is even, then the Fourier series of f reduces to the Fourier cosine series: $f_e(x) \sim a_0 + \sum_{n=1}^{\infty} a_n \cos nx$, with

$$a_0 = \frac{1}{L}\int_{0}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx$$
$$a_n = \frac{2}{L}\int_{0}^{L} f(x)\cos\left(\frac{n\pi x}{L}\right)dx, \quad n = 1,2,3,...$$

b. If f is odd, then the Fourier series reduces to the Fourier sine series:

$$f_o(x) \sim \sum_{n=1}^{\infty} b_n \sin nx, \text{ with } b_n = \frac{2}{L}\int_{0}^{L} f(x)\sin\left(\frac{n\pi x}{L}\right)dx, \quad n = 1,2,3,...$$

Example 2.2.1: consider the even function $f(x) = |x|$, $x \in [-1,1]$, and assume that f is periodic with period $p = 2L = 2$. The Fourier coefficients in the expansion of f are given by:

$$a_0 = \frac{1}{2}\int_{0}^{1} |x|\, dx = \frac{1}{2}. \quad \text{For } n \geq 1,$$
$$a_n = \frac{2}{1}\int_{0}^{1} |x|\cos(n\pi x)dx = 2\int_{0}^{1} x\cos(n\pi x)dx = \frac{2}{n^2\pi^2}(\cos n\pi - 1)$$

So,

$$a_n = \begin{cases} 0 & \text{if } n \text{ even} \\ \dfrac{-4}{n^2\pi^2} & \text{if } n \text{ odd} \end{cases}$$

$$f(x) \sim \frac{1}{2} + \sum_{k=1}^{\infty} \frac{4}{(2k-1)^2\pi^2} \cos((2k-1)\pi x) .$$

Even and Odd Functions

Before looking at further examples of Fourier series it is useful to distinguish between two classes of functions for which the Euler–Fourier formulas can be simplified. These are even and odd functions, which are characterized geometrically by the property of symmetry with respect to the *y*-axis and the origin, respectively**.**

Analytically, *f* is an **even function** if its domain contains the point $-x$ whenever it contains the point *x*, and if $f(x) = f(-x)$ for each *x* in the domain of *f*. Similarly, *f* is an **odd function** if its domain contains $-x$ whenever it contains *x*, and if $f(-x) = -f(x)$ for each *x* in the domain of *f* .

Even and odd functions are particularly important in applications of Fourier series since their Fourier series have special forms, which occur frequently in physical problems.

Definition 2.2.5 [21]: Even periodic extension

Suppose f is defined on the interval $[0, L]$. The periodic even extension of *f* is defined as: $f_e(x) = \begin{cases} f(x) & \text{for } 0 \le x \le L \\ f(-x) & \text{for } -L \le x < 0 \end{cases}$ and $f_e(x+L) = f(x)$

Definition 2.2.6 [21]: Odd periodic extension

Let f be a function defined on the interval $[0,L]$. The periodic odd extension of f is defined as: $f_o(x)=\begin{cases} f(x) & \text{for } 0<x\le L \\ 0 & \text{for } x=0 \\ -f(-x) & \text{for } -L\le x<0 \end{cases}$ and

$f_o(x+L)=f(x)$.

Example 2.2.2: consider the function $f(x)=x^2+1$, $x\in[0,1]$, the periodic odd extension of f is defined as: $f_o(x)=\begin{cases} x^2+1 & \text{for } 0<x\le 1 \\ 0 & \text{for } x=0 \\ -x^2-1 & \text{for } -1\le x<0 \end{cases}$

The graphs of f and f_o are shown in Figures 1 and 2 respectively.

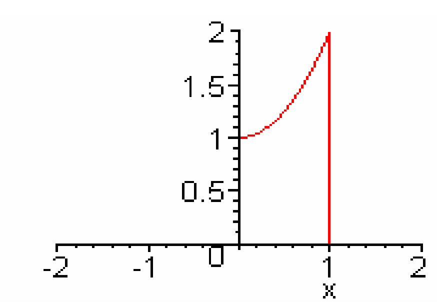


Figure 1

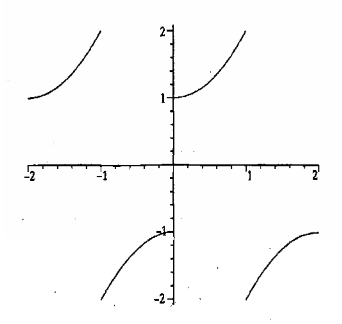


Figure 2

Example 2.2.3: let f be 2π periodic function defined on the interval $[-\pi,\pi]$, as

$$f(x)=\begin{cases} \pi-x\,, & x\in[0,\pi] \\ -\pi-x\,, & x\in[-\pi,0) \end{cases}$$

f is odd function so $a_n=0$ for $n\ge 0$, and $b_n=\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} f(x)\sin nx dx=\dfrac{2}{n}$.

So

$$f(x)\sim 2\sum_{n=1}^{\infty}\frac{\sin nx}{n}$$

Example 2.2.4: let $f(x)=\begin{cases} 0\,, & \text{if } x\in[-\pi,0) \\ 1\,, & \text{if } x\in[0,\pi) \end{cases}$

Then

$$f(x) \sim \frac{1}{2} + \sum_{n=1}^{\infty} \frac{2}{(2n-1)} \sin(2n-1)x .$$

2.3 Functional spaces

Definition 2.3.1: L^p-space

Let $p \geq 1$ be real number. Then the L^p-space is the set of all real-valued (or complex-valued) functions f on I, such that $\int_I |f(x)|^p \, dx < \infty$.

If $f \in L^p(I)$, then its L^p-norm defined as: $\|f\|_p = \left(\int_I |f(x)|^p \, dx \right)^{\frac{1}{p}}$.

Example 2.3.1:

a. The space $L^1(I)$ is the set of all integrable functions f on I, with L^1-norm defined by $\|f\|_1 = \int_I |f(x)| \, dx < \infty$.

b. The space $L^2(I)$ is the set of all square integrable functions f on I, with L^2 – norm defined by $\|f\|_2 = \left(\int_I |f(x)|^2 \, dx \right)^{\frac{1}{2}} < \infty$, and we say that the function has finite energy.

Remarks [1]:

a. Any continuous or piecewise continuous function with finite number of jump discontinuities on a finite closed interval I is in $L^1(I)$.

b. Any function bounded on finite interval I is square integrable on I. This includes continuous and piecewise continuous functions with finite jump discontinuities on a finite closed interval.

Theorem 2.3.2 [1]: Let I be a finite interval. If $f \in L^2(I)$, then $f \in L^1(I)$. In other words, a square integrable function on a finite interval is integrable.

Remarks [1]:

a. The conclusion of theorem 2.3.2 doesn't hold if I is an infinite interval, for example

$$f(x)=\begin{cases}\dfrac{1}{x}\ , & x\geq 1\\ 0\ , & x<1\end{cases}$$

$f \in L^2(I)$ but $f \notin L^1(I)$.

b. The converse of theorem 2.3.2 is not true, for example $f(x)=\dfrac{1}{\sqrt{x}}$, $x\in(0,1)$, is in $L^1((0,1))$ but not in $L^2((0,1))$.

Definition 2.3.3 [4]: The L^2-inner product on $L^2(I)$ is defined as $\langle f,g\rangle_{L^2} = \int_I f(x)\overline{g(x)}\,dx$, $f,g\in L^2(I)$, where $\overline{g}$ is the complex conjugate of g.

In case where the signal is discrete, we represent the signal as a sequence $X=\{x_n\}_{n=-\infty}^{\infty}$, where each x_n is the numerical value of the signal at the n^{th} time interval $[t_n, t_{n+1}]$.

Definition 2.3.4 [4]: Let $p\geq 1$ be real number. Then the l^p-space is the set of all real-valued (or complex-valued) sequences *X*, such that $\sum_{n=-\infty}^{\infty}|x_n|^p<\infty$.

The space l^2 is the set of all sequences X, with $\sum_{n=-\infty}^{\infty}|x_n|^2<\infty$. The inner product on this space is defined by

$$<X,Y>_{l^2}=\sum_{n=-\infty}^{\infty}x_n\overline{y_n}\ ,$$

where $X=\{x_n\}_{n=-\infty}^{\infty}$, and $Y=\{y_n\}_{n=-\infty}^{\infty}$.

Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of real-valued or complex-valued functions defined on some interval I of the real line. We consider four types of convergence:

a. **Pointwise convergence.** A sequence of functions f_n converges to f pointwise on I if for each $x \in I$ and for each small $\varepsilon > 0$, there exist a positive integer N such that if $n \geq N$, then $|f_n(x) - f(x)| < \varepsilon$.

b. **Uniform convergence.** A sequence of functions f_n converges to f uniformly on the interval I if for each small $\varepsilon > 0$, there exist a positive integer N such that if $n \geq N$, then $|f_n(x) - f(x)| < \varepsilon$.

c. **Convergence in L^2 – norm.** A sequence of functions f_n converges to f in $L^2(I)$ if $\|f_n(x) - f(x)\|_2 \to 0$ as $n \to \infty$, i.e given any $\varepsilon > 0$, there exist $N > 0$ such that if $n \geq N$, then $\|f_n(x) - f(x)\|_2 < \varepsilon$.

d. **Convergence in L^1 – norm**. A sequence of functions f_n converges to f in $L^1(I)$ if for any $\varepsilon > 0$, there exist $N > 0$ such that if $n \geq N$, $\|f_n(x) - f(x)\|_1 < \varepsilon$.

Remarks:

a. If the interval I is bounded, then the uniform convergence implies convergence in both L^1 and L^2 norm.

b. The uniform converge always implies the pointwise converges, but the converse is not true.

c. The uniform convergence is very useful when we want to approximate some function by sequence of continuous function $f_n(x)$.

Theorem 2.3.5: Uniform convergence theorem

Let $\{f_n\}_{n=-\infty}^{\infty}$ be a sequence of continuous functions on I and suppose $f_n \to f$ uniformly on I, then f is continuous function on I.

Proof: Suppose $f_n \to f$ uniformly and each f_n is continuous. Then given any $\varepsilon > 0$, there exist N such that $n > N$ implies $|f(x) - f_n(x)| < \frac{\varepsilon}{3}$ for all x.

Pick an arbitrary n larger than N. Since f_n is continuous, given any point $x_0 \in I$, $\exists \delta > 0$ such that $0 < |x - x_0| < \delta \Rightarrow |f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}$.

Therefore, given any $\varepsilon > 0$, $\exists \delta > 0$ such that $0 < |x - x_0| < \delta \Rightarrow$

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon . \end{aligned}$$

Therefore, f is continuous function on I.

2.4 Convergence of Fourier series

We start this section by discussing two important properties of the Fourier coefficients: Bessel's inequality which relates the energy of a square integrable function to its Fourier coefficients, and the Riemann–Lebesgue lemma ensures the vanishing of the Fourier coefficients of a function.

Theorem 2.4.1: (Bessel's inequality).

If f is a square integrable function on $[-\pi, \pi]$, i.e. $\int_{-\pi}^{\pi} |f(x)|^2 \, dx$ is finite, then

$$2|a_0|^2 + \sum_{n=1}^{\infty} |a_n|^2 + |b_n|^2 \leq \frac{1}{\pi} \int_{-\pi}^{\pi} |f(x)|^2$$

Where a_0, a_n, b_n are the Fourier coefficients of f.

Bessel's inequality says that if f has finite energy, then the module-square of the

Fourier coefficients are also finite.

Lemma 2.4.2 [4]: (The Riemann-Lebesgue Lemma)

Suppose f is piecewise continuous function on the interval $[a, b]$, Then

$$\lim_{n \to \infty} \int_a^b f(x) \sin nx \, dx = \lim_{n \to \infty} \int_a^b f(x) \cos nx \, dx = 0$$

Proof: consider the integral

$$\int_a^b f(x)\sin nx\,dx,$$

we have

$$\int_a^b f(x)\sin nx\,dx = \left.\frac{-f(x)\cos nx}{n}\right|_a^b + \int_a^b f'\frac{\cos nx}{n}dx$$

as $n \to \infty$, the right integral becomes zero (by using the sandwich theorem). So that

$$\lim_{n\to\infty}\int_a^b f(x)\sin nx\,dx = 0$$

Similarly,

$$\lim_{n\to\infty}\int_a^b f(x)\cos nx\,dx = 0.$$

There are two consequence of this theorem one of them is that only the first few terms in the Fourier series are the most important since they contribute more to the sum which means that only finite number of terms can be used to approximate the function. This is especially important in data compression. Another one is used to proof our convergence result.

Convergence theorems are concerned with how the partial sum

$$S_N(x) = a_0 + \sum_{n=1}^{N} a_n\cos nx + b_n\sin nx$$

converge to $f(x)$. The partial sum can be written in terms of an integral as follows:

$$\begin{aligned} S_N(x) &= a_0 + \sum_{n=1}^{N} a_n\cos nx + b_n\sin nx \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)dt + \frac{1}{\pi}\sum_{n=1}^{N}\left(\int_{-\pi}^{\pi} f(t)\cos(nx)\cos(nt)dt + \int_{-\pi}^{\pi} f(t)\sin(nx)\sin(nt)dt\right) . \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2} + \sum_{n=1}^{N}\cos(nx)\cos(nt) + \sin(nx)\sin(nt)\right)dt . \end{aligned}$$

$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2} + \sum_{n=1}^{N}\cos\big(n(t-x)\big)\right) dt\,.$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{\sin\big((N+1/2)(t-x)\big)}{\sin\big((t-x)/2\big)}\right) dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)D_N(t-x)\, dt$$

So, by change of variable $(u = t - x)$, and using lemma 2.2.2, we have

$$S_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(u+x)D_N(u)du\,,$$

where $D_N(u) = \dfrac{\sin(N+1/2)u}{2\sin(u/2)}$, is called **Dirichlet Kernel** of order *N*.

Convergence of Fourier series depends on the Dirichlet kernel. The following theorem states the basic property of this kernel.

Theorem 2.4.3 [19]: The Dirichlet kernel satisfies the following property:

a. Each $D_N(t)$ is real valued, continuous, 2π -periodic function

b. Each $D_N(t)$ is an even function.

c. For each N, $D_N(0) = N + \frac{1}{2}$, and $|D_N(t)| \le N + \frac{1}{2}$.

d. For each N, $\frac{1}{\pi}\int_I D_N(t)\, dt = \frac{2}{\pi}\int_0^{\pi} D_N(t)\, dt = 1$.

e. For $0 < |t| < \pi$, $|D_N(t)| \le \frac{\pi}{2|t|}$.

f. $2\|D_N(t)\|_I \to \infty,$ as $N \to \infty$.

Some of the features of the Dirichlet kernel can be seen Figure 3. The symmetry is certainly apparent ($D_N(t)$ is even) and that the graph oscillates above and below the horizontal axis is evident. The value of the function is small except close to 0 where the function is large, and as *N* increases this feature becomes more clear. The total area remains fixed always at π because of cancellations.

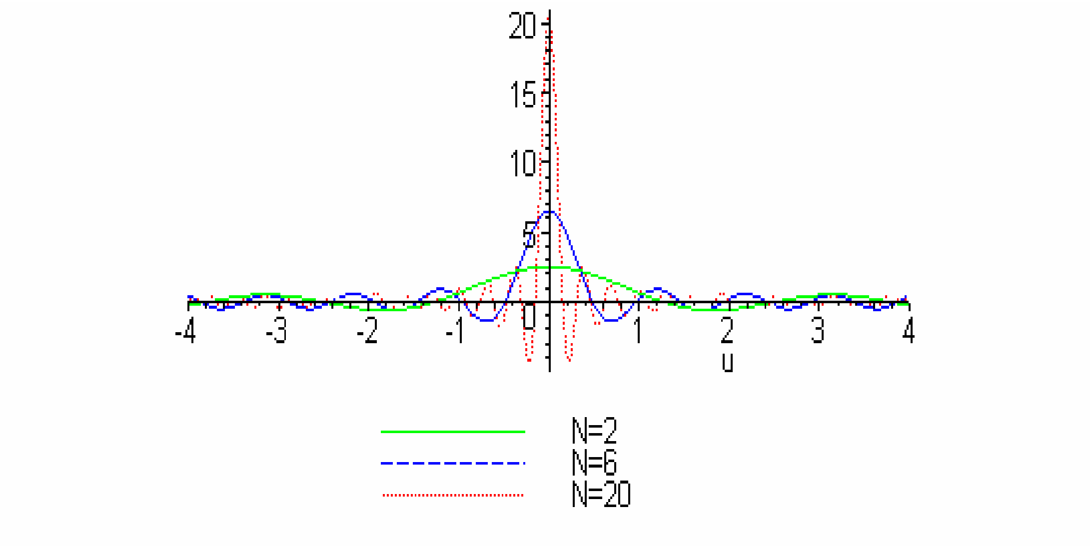


Figure 3 : $D_N(t)$

The following theorem gives conditions for convergence at a point of continuity.

Theorem 2.4.4 [4]: Suppose f is a continuous, 2π periodic function. Then for each point x where the derivative of f is defined, the Fourier series of f at x converges to $f(x)$.

Proof: let $S_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(u+x)D_N(u)\,du$, we want to show that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(u+x)D_N(u)\,du \to f(x) \text{ as } N \to \infty,$$

(by theorem 2.4.3, d) we have

$$f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)D_N(u)\,du,$$

so we must show that:

$$\frac{1}{\pi}\int_{-\pi}^{\pi}\left(f(u+x) - f(x)\right)D_N(u)\,du = \frac{1}{\pi}\int_{-\pi}^{\pi}\left(\frac{f(u+x)-f(x)}{\sin(u/2)}\right)\sin\left(\left(N+1/2\right)u\right)du \to 0$$

as $N \to \infty$.

Let $g(u)=\dfrac{f(u+x)-f(x)}{\sin(u/2)}$. The only possible value of $u\in[-\pi,\pi]$, where $g(u)$ could be discontinuous is $u=0$, so

$$\lim_{u\to 0} g(u)=\lim_{u\to 0}\frac{f(u+x)-f(x)}{u}.\frac{u/2}{\sin(u/2)}.2=f'(x).2.1=2f'(x).$$

Since f' is exist, then $g(u)$ is continuous and by Riemann- lebesgue lemma the last integral is zero as N large enough and this finish the proof.

Note that the hypothesis of this theorem requires the function f to be continuous. However, there are many functions of interest that are not continuous. So the following theorem gives conditions for convergence at a point of discontinuity.

Theorem 2.4.5 [4]: Suppose f is periodic and piecewise continuous, suppose x is a point where f is left and right differentiable (but not necessarily continuous).Then the Fourier series of f at x converge to $\dfrac{f(x+0)+f(x-0)}{2}$.

Proof: we must show that

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(u+x)D_N(u)du\to\frac{f(x+0)+f(x-0)}{2}\quad\text{as}\quad N\to\infty$$

where $\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{\pi} D_N(u)du=1$, in other words,

$$\frac{1}{\pi}\int_{0}^{\pi} f(u+x)D_N(u)du\to\frac{f(x+0)}{2}$$

$$\frac{1}{\pi}\int_{-\pi}^{0} f(u+x)D_N(u)du\to\frac{f(x-0)}{2}$$

these limits are equivalent to the following limits respectively,

$\dfrac{1}{\pi}\displaystyle\int_{0}^{\pi}(f(u+x)-f(x+0))D_N(u)du\to 0$, and $\dfrac{1}{\pi}\displaystyle\int_{-\pi}^{0}(f(u+x)-f(x-0))D_N(u)du\to 0$

by definition of $D_N(u)$ and Riemann lebesgue lemma we have

$$\frac{1}{2\pi}\int_0^{\pi}\left(\frac{f(u+x)-f(x+0)}{\sin(u/2)}\right)\sin((N+1/2)u)\,du \to 0$$

Let $g(u)=\dfrac{f(u+x)-f(x+0)}{\sin(u/2)}$,

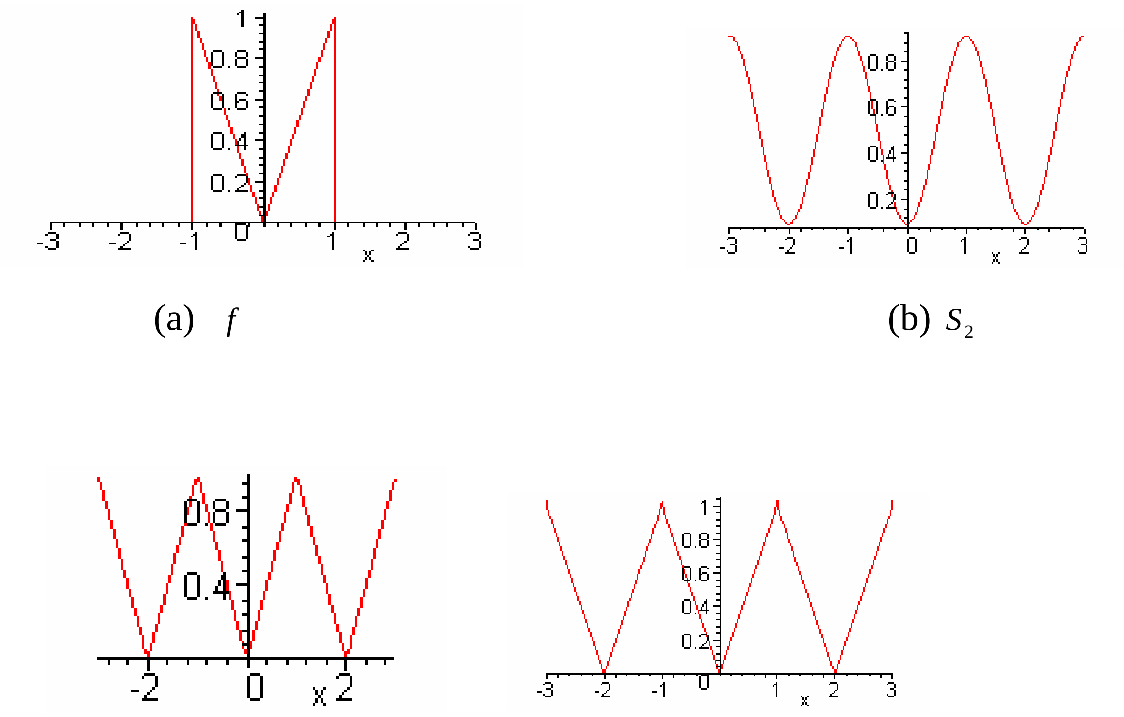
since u is positive its enough to show that $g(u)$ is continuous from the right

$$\lim_{u\to 0^+} g(u) = \lim_{u\to 0^+}\frac{f(u+x)-f(x+0)}{u}.\frac{u/2}{\sin(u/2)}.2 = f'(x+0).2.1 = 2f'(x+0).$$

since f is assumed to be right differentiable then the proof is finish.

Similarly, we can show that $\frac{1}{\pi}\int_{-\pi}^{0}(f(u+x)-f(x-0))D_N(u)du \to 0$ as $N\to\infty$.

In example (2.2.1), the function f is continuous on $[-1,1]$. Therefore, its Fourier series converges for all $x\in[-1,1]$. Figure 4 shows the graphs f together with the partial sums S_2, S_{10}, and S_{50} of its Fourier series.

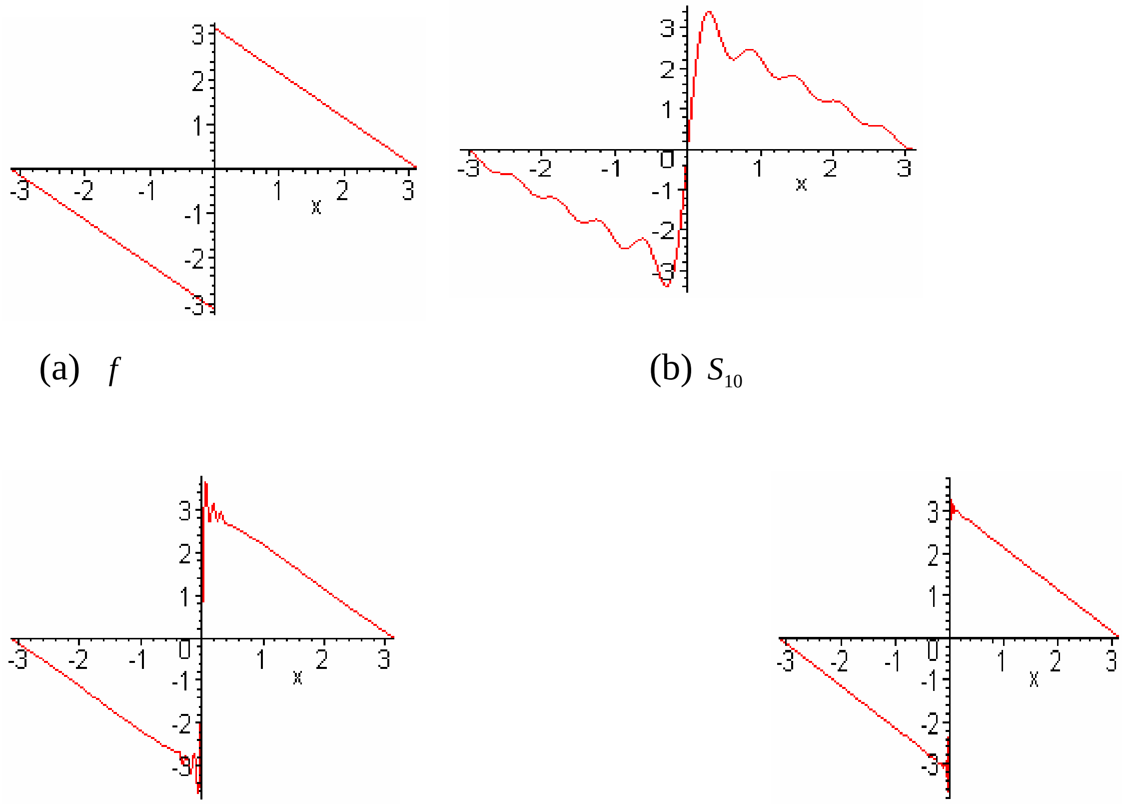


(a) f (b) S_2

(c) S_{10} (d) S_{50}

Figure 4

In example (2.2.3) f has a jump discontinuity at $x=0$, so Fourier series converges at each point in $[-\pi,\pi]$ except at $x=0$. Figure 5 shows the graphs f together with the partial sums S_{10}, S_{50} and S_{200} of its Fourier series.



(a) f (b) S_{10}

(c) S_{50} (d) S_{200}

Figure 5

The following theorem gives conditions for the uniform convergence.

Theorem 2.4.6 [4]: The Fourier series of piecewise smooth, 2π-periodic function f converges uniformly to f on $[-\pi,\pi]$.

Proof: To simplify the proof we can assume that the function f is twice differentiable. Consider the Fourier series of both f, f'', respectively;

$$f(x) \sim a_0 + \sum_{n=1}^{\infty} a_n \cos nx + b_n \sin nx,$$

$$f''(x) \sim a_0'' + \sum_{n=1}^{\infty} a_n'' \cos nx + b_n'' \sin nx,$$

we have the following relation between the coefficients of f and the coefficients of f'':

$$a_n = \frac{-1}{n^2} a_n''$$

$$b_n = \frac{-1}{n^2} b_n'' \ .$$

If f'' is continuous, then both the a_n'' and b_n'' stay bounded by some number M (in fact, by Riemann-Lebesgue lemma, a_n'' and b_n'' converges to zero as $n \to \infty$ .Therefore,

$$\sum_{n=1}^{\infty} |a_n| + |b_n| = \sum_{n=1}^{\infty} \frac{|a_n''| + |b_n''|}{n^2} \le \sum_{n=1}^{\infty} \frac{M + M}{n^2},$$

the last series is convergence and hence, $\sum_{n=1}^{\infty} \left(|a_n| + |b_n| \right) < \infty$.

$\left| f(x) - S_N(x) \right| = \left| \sum_{n=N+1}^{\infty} a_n \cos nx + b_n \sin nx \right| \le \sum_{n=N+1}^{\infty} \left(|a_n| + |b_n| \right)$ uniformly for all x .

But $\sum_{n=N+1}^{\infty} \left(|a_n| + |b_n| \right)$ is small for large N , so given $\varepsilon > 0, \exists N_0 > 0$ such that if $N > N_0$, then $\left| f(x) - S_N(x) \right| < \varepsilon, \forall x$. N doesn't depend on x, thus the convergence of $S_N(x)$ is uniformly.

Example 2.4.1: Gibbs phenomena [17]

Let's return to our example 2.2.3. f has a discontinuity at $x = 0$ so the convergence of its Fourier series can't be uniform. Let's examine this case carefully. What happens to the partial sums near the discontinuity?

Here, $S_N(x) = 2\sum_{n=1}^{N} \frac{\sin nx}{n}$ so

$$S_N'(x) = 2\sum_{n=1}^{N} \cos nx = \frac{\sin(N + 1/2)x}{\sin(x/2)} - 1 = 2\frac{\sin\frac{Nx}{2}\cos\frac{(N+1)x}{2}}{\sin\frac{x}{2}}, x \ne 0 .$$

Thus, since $S_N(0) = 0$ and we have

$$S_N(x) = \int_0^x S_N'(t)dt = \int_0^x 2\frac{\sin\frac{Nt}{2}\cos\frac{(N+1)t}{2}}{\sin\frac{t}{2}}dt \,.$$

Note that $S_N'(0) > 0$ so that S_N starts out at zero for $x = 0$ and then increases. Looking at the derivative of S_N we see that the first maximum is at the critical point $x_N = \frac{\pi}{N+1}$(the first zero of $\cos\frac{(N+1)x}{2}$ as x increases from 0). Here, $f(x_N) = \pi - x_N$.

The error is

$$S_N(x_N) - f(x_N) = \int_0^{x_N} 2\frac{\sin\left(N+\frac{1}{2}\right)t}{\sin\frac{t}{2}}dt - \pi \,.$$

$$= \int_0^{x_N} 2\frac{\sin\left(N+\frac{1}{2}\right)t}{t}dt + \int_0^{x_N}\left(\frac{2}{\sin(t/2)} - \frac{2}{t}\right).\sin\left(N+\frac{1}{2}\right)t\, dt - \pi \,.$$

$$= I(x_N) + J(x_N) - \pi \,.$$

Where

$$I(x_N) = \int_0^{x_N} 2\frac{\sin\left(N+\frac{1}{2}\right)t}{t}dt = \int_0^{(N+1/2)x_N} 2\frac{\sin u}{u}du \to \int_0^{\pi} 2\frac{\sin u}{u}du \approx 3.702794104$$

$$J(x_N) = \int_0^{x_N}\left(\frac{2}{\sin(t/2)} - \frac{2}{t}\right).[\sin Nt\cos\frac{t}{2} + \cos Nt\sin\frac{t}{2}]dt \,.$$

By Riemann-Lebesgue lemma $J(x_N) \to 0$ as $N \to \infty$.

We conclude that

$$\lim_{N\to\infty}[S_N(x_N) - f(x_N)] \approx 3.702794104 - \pi \approx .559.$$

The partial sum is overshooting the correct value by about 17.8635%! This is due the Gibbs Phenomenon. At the location of the discontinuity itself, the partial Fourier series will converge to the midpoint of the jump.

In mathematics, the Gibbs phenomenon, named after the American physicist J.Willard Gibbs, is the peculiar manner in which the Fourier series of a piecewise continuously differentiable periodic function f behaves at a jump discontinuity: the *n*th partial sum of the Fourier series has large oscillations near the jump, which might increase the maximum of the partial sum above that of the function itself. The overshoot does not die out as the frequency increases, but approaches a finite limit.

Note that the differentiability condition cited in theorems 2.4.4 through theorem 2.4.6 is to ensure the convergence of the Fourier series of f . So, in the case where the function is continuous but not piecewise differentiable, it's impossible to say that the Fourier series of such function is converge to f (pointwise or uniformly).

In 1873, **Due Bois-Raymond,** showed that there is a continuous function whose Fourier series diverge everywhere on accountably infinite set of point. The construction of this example is in [20]. Many years earlier **Kolmogorove** [5],(1926), had proved the existence of an example of a 2π - periodic, L^1 function that has Fourier series diverges at every points.

Kolmogorove example [5]: let $\{f_n\}_{n=1}^{\infty}$ be a sequence of trigonometrical polynomials of orders $\nu_1, \nu_2, \nu_3, \ldots$ with the following properties:

a. $f_n(x) \geq 0$.

b. $\int_0^{2\pi} f_n(x)dx = 2\pi$.

Moreover, suppose that to every f_n corresponds an integer λ_n, where $0 < \lambda_n \leq \nu_n$,a number A_n>0, and a point set E_n, such that

a. If $x \in E_n$, there is an integer $K = K_x$, $\lambda_n \le K \le \nu_n$ for which $S_K(x; f_n) > A_n$.

b. $A_n \to \infty$.

c. $\lambda_n \to \infty$.

d. $E_1 \subset E_2 \subset ... \subset,$, $E_1 + E_2 + ... = (0,2\pi)$.

Under these conditions, $\{n_K\}$ tends to ∞ sufficiently rapidly, the Fourier series of the function $f(x) = \sum_{K=1}^{\infty} \frac{f_{n_K}(x)}{\sqrt{A_{n_K}}}$, diverges every where.

The proof is very difficult, so you can found it in [5].

In the case where a Fourier series doesn't converge uniformly or pointwise it may be converge in weaker sense such as in L^2.i.e. Convergence in the mean

Theorem 2.4.7 [4]: suppose $f \in L^2([-\pi,\pi])$, let $f_N(x) = a_0 + \sum_{n=1}^{N} a_n \cos nx + b_n \sin nx$.

Where a_n, and b_n, n =0,1,2,…, are the coefficients of f, then f_N converge to f in L^2. i.e $\|f_N - f\|_2 \to 0$ as $N \to \infty$

Remark: f_N in V_n = the linear span of $\{1, \cos nx, \sin nx\}$, which is the closest in the L^2-norm, i.e. $\|f_N - f\|_2 = \min_{g \in V_n} \|g - f\|_2$

*Proof***:** The proof consists of two steps:

1^{st} step, any function can be approximated arbitrarily by a smooth, 2π-periodic function say *g*.

2^{nd} step, this function *g* can be approximate uniformly and therefore in L^2 by its Fourier series.

Assume we proved the 1^{st} step, so for any $f \in L^2([-\pi,\pi])$, there exists a 2π-periodic and smooth function g such that:

$$\|g-f\|_2 < \varepsilon \qquad (2.4.1)$$

Let $g_N(x) = c_0 + \sum_{n=1}^{N} c_n \cos nx + d_n \sin nx$, where c_n, d_n are the coefficients of g.

Since g is differentiable, then we can approximate g uniformly by g_N, by choosing N_0 large enough such that

$$|g(x) - g_N(x)| < \varepsilon, \forall x \in [-\pi, \pi] \qquad (2.4.2)$$

for $N > N_0$, we have

$$\|g - g_N\|^2 = \int_{-\pi}^{\pi} |g(x) - g_N(x)|^2 dx \le \int_{-\pi}^{\pi} \varepsilon^2 dx = 2\pi\varepsilon^2 \qquad (2.4.3)$$

$$\Rightarrow \|g - g_N\| < \sqrt{2\pi}\varepsilon \qquad (2.4.4)$$

by (2.4.1)$\rightarrow$ (2.4.4)

$$\|f - g_N\|_2 = \|f - g + g - g_N\| \le \|g - f\| + \|g - g_N\|$$

$$< \varepsilon + \sqrt{2\pi}\varepsilon, \text{ for } N > N_0,$$

but g_N in V_n, so

$$\|f_N - f\|_2 = \min_{g \in V_n} \|g - f\|_2 < \|f - g_N\| < \left(1 + \sqrt{2\pi}\right)\varepsilon, \text{ for } N > N_0$$

since ε arbitrary $\Rightarrow$ the proof is finish.

2.5 Summability of Fourier series

A study of convergence property of Fourier series partial sum will face some problems, such as Kolomogrove example, and Gibb's phenomenon in the partial sums for discontinuous function, finally, Du' Bois Raymond example of continuous function whose Fourier series diverge some where. All of these difficulties can be solved by using other summation formula or method, one of them is to take the arithmetic mean of the partial sums of the Fourier series [19]:

$$\sigma_N(x) = \left(S_0(x) + S_1(x) + \ldots\ldots + S_{N-1}(x)\right)/N . \qquad (2.5.1)$$

$$= \frac{1}{N\pi}\sum_{j=0}^{N-1}\int_0^{2\pi} D_j(x-t)f(t)dt$$

$$= \int_0^{2\pi}\left[\frac{1}{N\pi}\sum_{j=0}^{N-1} D_j(x-t)\right]f(t)dt$$

$$= \frac{1}{\pi}\int_0^{2\pi} K_N(x-t)\, f(t)dt$$

where

$K_N(x) = \frac{1}{N}\sum_{j=0}^{N-1} D_j(x) = \frac{1}{N}\left(\frac{\sin Nx/2}{\sin x/2}\right)^2$, is called **Fejer Kernel** of order N.

The idea of forming averages for divergent series formula studied by Ernesto Cesaro [19] in 1890, and then the mathematician Leopold Fejer [19] first applied it in 1990 to study the Fourier series and he had shown that Cesaro summability was a way to overcome the problem of divergence of a classical Fourier series for the case of continuous functions.

Now, we will set the basic properties of this kernel in the following theorem

Theorem 2.5.1 [19]: (Properties of Fejer kernel)

Let $K_N(x)$ be the Fejer Kernel.

a. Each $K_N(x)$ is real valued, non negative, continuous function.

b. Each $K_N(x)$ is an even function.

c. For each N, $\frac{1}{\pi}\int_I K_N(x)dx = \frac{2}{\pi}\int_0^{\pi} K_N(x)dx = 1$.

d. For each $N, K_N(0) = N$.

The reason why the formula (2.5.1) is better properties than ordinary partial sums is that the Fejer kernel is nonnegative. So, its graph here doesn't oscillate above and below the horizontal axis like Diriklet kernel, but remains on or above. The total area under the graph of Fejer kernel (see

Figure 6) remains fixed at π, but this is not because of any cancellation, and for this reason the Cesaro means of the Fourier series of continuous function can converge even though the series diverges.

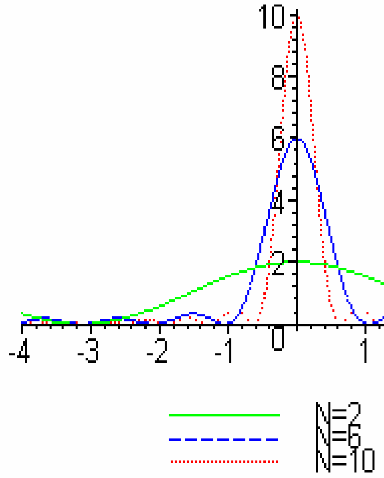


Figure 6

The following theorem gives conditions for the Convergence in Cesaro mean.

Theorem 2.5.2 [19]: let f be integrable function, and let $\sigma_N(x)$ denote the Cesaro mean of the Fourier series of f, if f is piecewise continuous, x_0 is the point of discontinuity, then

$$\lim_{N\to\infty} \sigma_N(x) = \frac{f(x_0+0)+f(x_0-0)}{2},$$

Moreover, If f is a 2π-periodic function that is continuous at each point on *I*, then $\sigma_N(x)$ converge to f uniformly for each x in *I*.

Proof: let $\varepsilon > 0$ choose $\delta > 0$ such that for every $0 \le t \le \delta$, we have

$$\left| f(x_0+t) + f(x_0-t) - 2f(x_0) \right| < \varepsilon \qquad (2.5.2)$$

By theorem (2.5.1, c) the integral $\frac{2}{\pi}\int_0^{\pi} f(x_0)K_N(t)dt = f(x_0)$,

$$|\sigma_N(x) - f(x)| = \left|\frac{1}{\pi}\int_0^{\pi}\big(f(x_0+t)+f(x_0-t)\big)K_N(t)dt - \frac{2}{\pi}\int_0^{\pi} f(x_0)K_N(t)dt\right|$$

$$= \left|\frac{1}{\pi}\int_0^{\pi}\big(f(x_0+t)+f(x_0-t)-2f(x_0)\big)K_N(t)dt\right|$$

$$\leq \frac{1}{\pi}\int_0^{\pi}\left|f(x_0+t)+f(x_0-t)-2f(x_0)\right|K_N(t)dt$$

$$= I_1 + I_2$$

where I_1 is the integral over the interval $[0,\delta]$, and I_2 is the integral over the interval $[\delta,\pi]$.

By (2.5.2), $I_1 \leq \frac{\varepsilon}{\pi}\int_0^{\delta} K_N(t)dt \leq \varepsilon$, and for large N , I_1 becomes small, because the bound of the size of $K_N(t)$ for t away from zero.

Let $\kappa_N = \sup\{K_N(t), \delta \leq t \leq \pi\}$, by theorem (2.5.1, f) $\kappa_N \to 0$ as $N \to \infty$. So,

$$I_2 \leq \frac{\kappa_N \varepsilon}{\pi}\int_{\delta}^{\pi} |f(x_0+t)| + |f(x_0-t)| + 2|f(x_0)|dt \,.$$

So, for large N, I_2 becomes small, and since ε is arbitrary, then $\lim_{N\to\infty}\sigma_N(x_0) = f(x_0)$ and if f is continuous at each point on I, then the last limit apply uniformly. So that $\sigma_N(x)$ converge to f uniformly for each x in I.

Lemma 2.5.3 [17]: Suppose $f \in L^2([-\pi,\pi])$ and 2π-periodic function is bounded by M , then $|\sigma_N(x)| \leq M \quad \forall x$ and for all N .

As a result of lemma 2.5.3, Gibbs phenomenon will disappear. To show this, we use the sandwich theorem.

$$0 \leq \|f - \sigma_N\| \leq \|f\| - \|\sigma_N\|$$

$$0 \leq \lim_{N\to\infty}\|f - \sigma_N\| \leq \lim_{N\to\infty}\|f\| - \lim_{N\to\infty}\|\sigma_N\|$$

$$\leq M - M = 0$$

Hence, $\lim_{N\to\infty}\|f-\sigma_N\|=0$.

2.6 Generalized Fourier series

The classical theory of Fourier series has undergone extensive generalizations during the last two hundred years. For example, Fourier series can be viewed as one aspect of a general theory of *orthogonal series expansions*. Later, we shall discuss a few of the more orthogonal series, such as Haar series, and wavelet series. But now we give a formal definition of orthogonality of such system .

Definition 2.6.1 [1]: Orthogonality

A collection of functions $\{g_n(x)\}_{n\in\mathrm{N}} \in L^2(I)$ forms an orthogonal system on I if:

a. $\int_I g_n(x)\overline{g_m(x)}dx = 0$ for $n \neq m$.

b. $\int_I g_n(x)\overline{g_n(x)}dx = \int_I |g_n(x)|^2 dx > 0$

where $\overline{g}$ is the complex conjugate of g.

If in addition:

c. $\int_I g_n(x)\overline{g_n(x)}dx = \int_I |g_n(x)|^2 dx = 1$.

Then the system is orthonormal on I

Example 2.6.1:

The set $\{1, \sin(nx), \cos(nx)\}_{n\in\mathrm{N}}$ is an orthogonal system over $[-\pi, \pi]$, and the set $\left\{\frac{1}{\sqrt{2\pi}}, \frac{1}{\sqrt{\pi}}\sin(nx), \frac{1}{\sqrt{\pi}}\cos(nx)\right\}_{n\in\mathrm{N}}$ is an orthonormal system over the interval $[-\pi, \pi]$.

Definition 2.6.2 [1]: Generalized Fourier series

Let $f \in L^2(I)$ and let $\{g_n(x)\}_{n\in\mathrm{N}}$ be an orthonormal system on *I*. The generalized Fourier series is:

$$f(x) \sim \sum_{n\in\mathrm{N}} < f, g_n > g_n(x).$$

The fundamental question about Fourier series is: When is an arbitrary function equal to its Fourier series and in what sense does that Fourier series converge? The answer lies in the notation of a complete orthonormal system.

Definition 2.6.3 [1]: Given a collection of functions $\{g_n(x)\}_{n\in\mathrm{N}} \in L^2(I)$, the span of $\{g_n(x)\}_{n\in\mathrm{N}}$ denoted by $\mathrm{span}\{g_n(x)\}_{n\in\mathrm{N}}$ is the collection of all finite linear combinations of the elements of $\{g_n(x)\}_{n\in\mathrm{N}}$. The mean-square closure of $\mathrm{span}\{g_n(x)\}_{n\in\mathrm{N}}$, denoted $\overline{\mathrm{span}}(g_n(x))$ is defined as follows: A function $f \in \overline{\mathrm{span}}(g_n(x))$ if for every $\varepsilon > 0$, there is a function $g(x) \in \mathrm{span}\{g_n(x)\}_{n\in\mathrm{N}}$ such that $\|f - g\|_2 < \varepsilon$.

Definition: 2.6.4 [1]: Completeness

If every function in $L^2(I)$ is in $\overline{\mathrm{span}}(g_n(x))$ where $\{g_n(x)\}_{n\in\mathrm{N}}$ is orthonormal system, then we say that $\{g_n(x)\}_{n\in\mathrm{N}}$ is **complete** on *I*, this means that every function in $L^2(I)$ is equal to its Fourier series in $L^2(I)$. A complete orthonormal system is called an **orthonormal basis**.

The following two lemmas related to very important inequalities that will be very useful in the next theorem.

Lemma 2.6.5 [1]: Let $\{g_n(x)\}_{n\in \mathrm{N}}$ is the orthonormal system on *I*, then for every $f \in L^2(I)$,

$$\left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2^2 = \|f\|_2^2 - \sum_{n=1}^{N} |< f, g_n >|^2$$

The next theorem gives several equivalent criteria for an orthonormal system to be complete.

Lemma 2.6.6 [1]: Let $\{g_n(x)\}_{n\in \mathrm{N}}$ is the orthonormal system on *I*, then for every $f \in L^2(I)$, and every finite sequence of numbers $\{a(n)\}_{n=1}^{N}$

$$\left\| f - \sum_{n=1}^{N} a(n) g_n \right\|_2^2 = \left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2^2 + \sum_{n=1}^{N} |a(n) - < f, g_n >|^2 .$$

Theorem 2.6.7 [1]: Let $\{g_n(x)\}_{n\in \mathrm{N}}$ be an orthonormal system on *I* then the following are equivalent.

a. $\{g_n(x)\}_{n\in \mathrm{N}}$ is complete on *I*.

b. For every $f \in L^2(I)$, $f(x) = \sum_{n\in \mathrm{N}} < f, g_n > g_n(x)$ in $L^2(I)$.

c. Every function $f, C_c^{\,0}$ on *I* can be written as $f(x) = \sum_{n\in \mathrm{N}} < f, g_n > g_n(x)$, and $\|f\|_2^2 = \int_I |f(x)|^2 dx = \sum_{n\in \mathrm{N}} |< f, g_n >|^2$.

The last statement convert the inequality in *Bessel's inequality* to equality, which means that the sum of the moduli-squared of the Fourier coefficient is precisely the same as the energy of f .

Proof: $a \Rightarrow b$

If $\{g_n(x)\}_{n\in\mathbb{N}}$ is complete, by definition of a complete set, every $f \in L^2(I)$ is in $\overline{\text{span}}(g_n(x))$, so let $\varepsilon > 0$, then there exist a finite sequence $\{a(n)\}_{n=1}^{N_0}$, $N_0 \in \mathbb{N}$ (by definition of $\overline{\text{span}}(g_n(x))$), such that

$$\left\| f - \sum_{n=1}^{N} a(n) g_n \right\|_2 < \varepsilon .$$

So by lemma (2.6.5)

$$\left\| f - \sum_{n=1}^{N_0} < f, g_n > g_n \right\|_2^2 \le \left\| f - \sum_{n=1}^{N_0} < f, g_n > g_n \right\|_2^2 + \sum_{n=1}^{N_0} \left| a(n) - < f, g_n > \right|^2$$

$$= \left\| f - \sum_{n=1}^{N_0} a(n) g_n \right\|_2^2 < \varepsilon^2 .$$

But $\left\{ \left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2 \right\}_{n\in\mathbb{N}}$ is decreasing sequence, so for every $N \ge N_0$

$$\left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2^2 < \varepsilon .$$

$b \Rightarrow c$

Every function $f, C_c^{\,0}$ on I is in $L^2(I)$, by (b): $f(x) = \sum_{n\in\mathbb{N}} < f, g_n > g_n(x)$.

But the last equation hold iff $\lim_{N\to 0} \left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2^2 = 0$ for all $f, C_c^{\,0}$ on I.

by lemma (2.6.6), we have

$$\left\| f - \sum_{n=1}^{N} < f, g_n > g_n \right\|_2^2 = \| f \|_2^2 - \sum_{n=1}^{N} \left| < f, g_n > \right|^2$$

and this equivalent to $\lim_{N\to 0} (\| f \|_2^2 - \sum_{n=1}^{N} \left| < f, g_n > \right|^2) = 0$, hence c hold.

2.7 Fourier Transform

The Fourier transform can be thought of as a continuous form of Fourier series. A Fourier series decomposes a signal on $[-\pi, \pi]$ into components that vibrate at integer frequencies. By contrast, the Fourier transform decomposes a signal defined on an infinite time interval into a w-frequency component, where w can be any real (or even complex number).

As we have seen, any sufficiently smooth function f that is periodic can be built out of sine and cosine. We can also see that complex exponentials may be used in place of sine and cosine. We shall now use complex exponentials because they lead to less and simpler computations.

If f has period $2L$, its complex Fourier series expansion is

$$f(x)=\sum_{n=\infty}^{\infty}c_n e^{\frac{in\pi x}{L}} \quad , \quad \text{with}\, c_n=\frac{1}{2L}\int_{-L}^{L}f(x)e^{\frac{-in\pi x}{L}}dx\,.$$

Non-periodic functions can be considered as periodic functions with period $L=\infty$, and the Fourier series becomes Fourier integral

Fourier transform on $L^1(\Re)$

Definition 2.7.1 [12]: Fourier transform on $L^1(\Re)$

Let $f\in L^1(\Re)$, the Fourier transform of $f(x)$ of is denoted by $\hat{f}(w)$ and defined by

$$\hat{f}(w)=\int_{-\infty}^{\infty}f(x)e^{-iwx}dx$$

Physically, the Fourier transform, $\hat{f}(w)$, measures oscillation of $f(x)$ at the frequency w, and $\hat{f}(w)$ is called frequency spectrum of a signal or waveform $f(x)$.

Theorem 2.7.2 [4]: (Fourier inversion formula)

If $f\in L^1(\Re)$ is continuously differentiable function, then

$$f(x)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(w)e^{iwx}dw$$

If the function $f(x)$ has points of discontinuity, then the preceding formula holds with $f(x)$ replaced by the average of the left and right hand limits.

Note: The assumption $f \in L^1(\Re)$ in theorem (2.7.2) is needed to ensure that the improper integral defining $\hat{f}(w)$ converges.

Proof: we want to prove that $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(t)e^{-i(t-x)w}dtdw$

If f is non zero only finite interval, then the $t-$integral occurs only on this finite interval. The $w-$ integral still involves on infinite interval and this must be handled by integrating over a finite interval of the form $-L \le w \le L$, and then letting $L \to \infty$.

So we must show that $f(x) = \frac{1}{2\pi}\lim_{L\to\infty}\int_{-L}^{L}\int_{-\infty}^{\infty} f(t)e^{-i(t-x)w}dtdw$.

Using the definition of complex exponential $e^{iu} = \cos u + i\sin u$, the preceding limit is equivalent to showing

$$f(x) = \frac{1}{2\pi}\lim_{L\to\infty}\int_{-L}^{L}\int_{-\infty}^{\infty} f(t)[\cos(t-x)w - i\sin(t-x)w]dtdw.$$

Since sine is an odd function, the $w-$ integral involving $\sin(t-x)w = 0$, so

$$f(x) = \frac{1}{\pi}\lim_{L\to\infty}\int_{0}^{L}\int_{-\infty}^{\infty} f(t)[\cos(t-x)w]dtdw$$

and this is because cosine is an even function. now $\int_0^L \cos(t-x)wdw = \frac{\sin(t-x)L}{t-x}$, replacing t by $x+u$, the preceding limit is equivalent to

$$f(x) = \frac{1}{\pi}\lim_{L\to\infty}\int_{-\infty}^{\infty} f(x+u)\frac{\sin(Lu)}{u}du \tag{2.7.1}$$

To prove (2.7.1), we must show that for any $\varepsilon > 0$, the difference between $f(x)$ and the integral on the right is less than ε for sufficiently large *L*. For this ε, we can choose $\delta > 0$ such that

$$\frac{1}{\pi}\int_{-\delta}^{\delta}|f(x+u)|du < \varepsilon \tag{2.7.2}$$

we will use this inequality at the end of the proof.

Now we need to use the Riemann- Lebesgue lemma which state.

$\lim_{L\to\infty}\int_a^b g(u)\sin(Lu)du = 0$, where g is any piecewise continuous function. Here, a and b could be infinity if g is nonzero only on a finite interval. By letting $g(u) = f(x+u)/u$, we get the integrals

$$\frac{1}{\pi}\int_{-\infty}^{\delta} f(x+u)\frac{\sin(Lu)}{u}du \quad \text{and} \quad \frac{1}{\pi}\int_{\delta}^{\infty} f(x+u)\frac{\sin(Lu)}{u}du$$

which tends to zero as $L\to\infty$. Thus the limit in (2.7.1) is equivalent to showing

$$f(x) = \frac{1}{\pi}\lim_{L\to\infty}\int_{-\delta}^{\delta} f(x+u)\frac{\sin(Lu)}{u}du \tag{2.7.3}$$

but $f(x) = \frac{1}{\pi}\lim_{n\to\infty}\int_{-\pi}^{\pi} f(x+u)\frac{\sin((n+1/2)u)}{2\sin(u/2)}du$ (2.7.4)

(See theorem 2.4.4), so the proof of (2.7.3) will proceed in two steps.

Step 1:

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+u)\frac{\sin((n+1/2)u)}{2\sin(u/2)}du - \frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\frac{\sin((n+1/2)u)}{u}du$$

$$= \frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\sin((n+1/2)u)\left[\frac{1}{2\sin(u/2)} - \frac{1}{u}\right]du$$

since the integration over $(-\pi,-\delta)$ and (δ,π) is zero as $n\to\infty$, by Riemann-lebesgue lemma.

In addition, the quantity $\left(\frac{1}{2\sin(u/2)} - \frac{1}{u}\right)$ is continuous on the interval $-\delta \le u \le \delta$, because the only possible discontinuity occurs at $u = 0$, and the limit of this expression as $u \to 0$ is zero. So

$$\frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\sin((n+1/2)u)\left[\frac{1}{2\sin(u/2)} - \frac{1}{u}\right]du \to 0 \text{ as } n\to\infty.$$

Together with (2.7.4) , we show that

$$\frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\frac{\sin((n+1/2)u)}{u}du \to f(x) \text{ as } n\to\infty \qquad (2.7.5)$$

Which is the same limit in (2.7.3) for L of the form $L = n + 1/2$.

Step 2:

Any $L > 0$ can be written as $L = n + h$, $h \in [0,1)$, to show

$$\frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\left[\frac{\sin((n+1/2)u)}{u} - \frac{\sin Lu}{u}\right]du < \frac{\varepsilon}{2}$$

By using mean value theorem, we have

$$\left|\sin((n+1/2)u) - \sin Lu\right| = \left|\sin((n+1/2)u) - \sin(n+h)u\right|$$

$$= \left|\cos t\right|\left|u/2 - hu\right| \le \left|u\right|/2, \text{ since } h \in [0,1).$$

Therefore,

$$\frac{1}{\pi}\int_{-\delta}^{\delta}\left| f(x+u)\left(\frac{\sin((n+1/2)u) - \sin Lu}{u}\right)\right|du \le \frac{1}{\pi}\int_{-\delta}^{\delta}\left| f(x+u)\right|.\frac{|u|}{2|u|}du \le \frac{\varepsilon}{2}$$

Finally, we can choose N large enough so that if $n > N$, then

$$\left| f(x) - \frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\frac{\sin((n+1/2)u)}{u}du\right| < \frac{\varepsilon}{2}$$

this inequality together with the one in step (2.7.2) yields.

$$\left| f(x) - \frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\frac{\sin(Lu)}{u}du\right| <$$

$$\left| f(x) - \frac{1}{\pi}\int_{-\delta}^{\delta} f(x+u)\frac{\sin((n+1/2)u)}{u}du\right| + \frac{1}{\pi}\int_{-\delta}^{\delta}\left| f(x+u)\left(\frac{\sin((n+1/2)u) - \sin Lu}{u}\right)\right|du$$

$< \dfrac{\varepsilon}{2} + \dfrac{\varepsilon}{2} = \varepsilon,$ If $n > N$. Hence the proof is complete.

Example 2.7.1: The Fourier transform of $f(x) = \begin{cases} x+\pi, & x \in [-\pi, 0] \\ \pi - x, & x \in (0, \pi] \end{cases}$

Is given by

$$\hat{f}(w) = \sqrt{\frac{2}{\pi}}\frac{(1-\cos w\pi)}{w^2}$$

the graph of f and its Fourier transform are given in Figure (10).

(a): $f(x)$ (b): $\hat{f}(w)$

Figure 10

Example 2.7.2: Characteristic function

Let $\chi_\tau(x) = \begin{cases} 1, & x \in (-\tau, \tau) \\ 0 & \text{otherwise} \end{cases}$, then $\hat{\chi}_\tau(w) = \dfrac{2}{w} \sin w\tau$.

Note that $\chi_\tau(x) \in L^1(\Re)$, but its Fourier transform is not in $L^1(\Re)$. The graph of $\chi_\tau(x)$ and $\hat{\chi}_\tau(w)$ is given in Figure (11).

(a) : $\chi_\tau(x)$ (b) : $\hat{\chi}_\tau(w)$

Figure 11

Remarks [12]:

a. Note that the Fourier transform in example (2.7.1) decay at the rate $\dfrac{1}{w^2}$ as $w \to \infty$, which is faster than the decay rate of $\dfrac{1}{w}$ exhibited by the Fourier transform in example (2.7.2), the faster decay in example

(2.7.1) result from the continuity of the function. Note the similarity to the Fourier coefficients a_n , b_n in examples 2.2.1 and 2.2.3 of section 2.2.

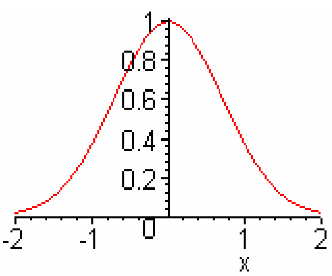
b. Some elementary functions, such as the constant function c , $\cos ax$, $\sin ax$, do not belong to $L^1(\Re)$, and hence do not have Fourier transform. But when these functions are multiplied by the characteristic function $\chi_\tau(x)$, the resulting functions belong to $L^1(\Re)$, and have Fourier transform.

Example 2.7.3: Gaussian function

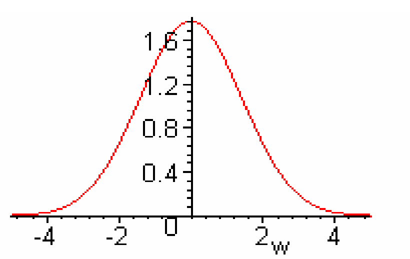
The Fourier transform of Gaussian function $f(x) = e^{-a^2x^2}$ is defined by

$\hat{f}(w) = \frac{\sqrt{\pi}}{4} e^{\frac{-w^2}{4a^2}}$, where $a > 0$.

The graph of $f(x)$, $\hat{f}(w)$ is given in Figure (12). Note that the Fourier transform of Gaussian function, is again Gaussian function.



(a): $f(t)$ at $a = 1$



(b): $\hat{f}(w)$ at $a = 1$

Figure 12

Basic Properties of Fourier transform

In this section, we set down most of the basic properties of the Fourier transform. First, we introduce the alternative notation $F(f)(w) = \hat{f}(w)$ for the Fourier transform of $f(x)$ and $F^{-1}(f)(x)$ for the inverse Fourier transform.

Theorem 2.7.3 [4]: Let f and g be differentiable functions defined on the real line with $f(x)=0$ for large $|x|$, then the following properties holds:

1. **Linearity:** The Fourier transform and its inverse are linear operator. That is for any constant c

   - $F(f+g)=F(f)+F(g)$ and $F(cf)=cF(f)$.

   - $F^{-1}(f+g)=F^{-1}(f)+F^{-1}(g)$ and $F^{-1}(cf)=cF^{-1}(f)$.

2. **Translation:** $F(f(x-a))(w)=e^{-iwa}F(f)(w)$.
3. **Rescaling:** $F(f(bx)(w)=\dfrac{1}{b}F(f)(\dfrac{w}{b})$.
4. The Fourier transform of a product of f with x^n is

   $F(x^n f(x))(w)=(i^n)\dfrac{d^n}{dw^n}\{F(f)(w)\}$.

5. The inverse Fourier transform of a product of f with w^n is

   $F^{-1}(w^n f(w))(x)=(-i^n)\dfrac{d^n}{dt^n}\{F^{-1}(f)(x)\}$

6. The Fourier transform of an n^{th} derivative is

   $F(f^{(n)}(x))(w)=(iw)^n\{F(f)(w)\}$

7. The inverse Fourier transform of n^{th} derivative is

   $F^{-1}(f^{(n)})(x)=(-ix)^n\{F^{-1}(f)(x)\}$.

Note that we assume that f is differentiable function with compact support , and we don't say that $f\in L^1(\Re)$, and this is because the Fourier transform of some function in $L^1(\Re)$ like the characteristic function, do not belong to the L^1- space, hence we can't talk about the inverse of the Fourier transform.

Theorem 2.7.4 [12]: Continuity

If $f\in L^1(\Re)$, then $\hat{f}(w)$ is continuous on $\Re$.

Proof : for any $w, h \in \Re$, we have

$$\left|\hat{f}(w+h)-\hat{f}(w)\right| = \left|\int_{-\infty}^{\infty} e^{-iwx}(e^{-ihx}-1)f(x)dx\right|$$
$$\leq \int_{-\infty}^{\infty} \left|e^{-ihx}-1\right|\left|f(x)\right|dx$$

since $\left|e^{-ihx}-1\right|\left|f(x)\right| \leq 2.\left|f(x)\right|$ and $\lim_{h\to 0}\left|e^{-ihx}-1\right| = 0, \forall x \in \Re$

we conclude that as $h \to 0$, $\left|\hat{f}(w+h)-\hat{f}(w)\right| \to 0$.

Which is independent of w, by the lebesgue dominated convergence theorem. This proves that $\hat{f}(w)$ is continuous on $\Re$. In fact, $\hat{f}(w)$ is uniformly continuous on $\Re$.

Theorem 2.7.5 [12]: (Riemann- Lebesgue lemma)

If $f \in L^1(\Re)$, then $\lim_{|w|\to\infty}\left|\hat{f}(w)\right| = 0$

Proof : since $e^{-iwx} = -e^{-iw(x+\pi/w)}$, we have

$$\hat{f}(w) = -\int_{-\infty}^{\infty} f(x-\frac{\pi}{w})e^{-iwx}\ dx = -\int_{-\infty}^{\infty} f(x)e^{-iw(x+\pi/w)}dx,$$

Thus,

$$\hat{f}(w) = \frac{1}{2}\left[\int_{-\infty}^{\infty} f(x)e^{-iwx}\ dx - \int_{-\infty}^{\infty} f(x-\frac{\pi}{w})e^{-iwx}\ dx\right]$$
$$= \frac{1}{2}\int_{-\infty}^{\infty}\left[f(x)-f(x-\frac{\pi}{w})\right]e^{-iwx}\ dx$$

clearly,

$$\lim_{|w|\to\infty}\left|\hat{f}(w)\right| \leq \frac{1}{2}\lim_{|w|\to\infty}\int_{-\infty}^{\infty}\left|f(x)-f(x-\frac{\pi}{w})e^{-iwx}\right|dx = 0$$

Observe that the space $C_0(\Re)$ of all continuous on $\Re$ which decay at infinity, that is $f(x) \to 0$ as $|x| \to \infty$, is norm space with respect to the norm defined by $\|f\| = \operatorname{Sup}_{x\in\Re}|f|$.

It follows from above theorem that the Fourier transform is continuous linear operator from $L^1(\Re)$ to $C_0(\Re)$.

Fourier transform on $L^2(\Re)$

Until now, we have been making the assumption that a function f must be in $L^1(\Re)$ in order for its Fourier transform to be defined. But we have seen example like the constant function doesn't belong to $L^1(\Re)$, suggest that we need to expand the definition to a large class of functions, L^2- functions. The formal definition (2.7.1) of the Fourier transform doesn’t make sense for a general $f \in L^2(\Re)$, because there is a square integrable function do not belong to $L^1(\Re)$, and hence $\hat{f}(w)$ doesn’t converge . So, we can define the Fourier transform for such function as follows:

Let $f \in L^2(\Re)$, then $f_N = f\,\chi_{[-N,N]} \in L^1(\Re)$, now the space of step functions is dense in $L^2(\Re)$**,** so we can fined a convergent sequence of step functions $\{s_n\}$ such that $\lim\limits_{n\to\infty}\|f - s_n\|_{L^2} = 0$.

Note that the sequence of functions $\{f_N = f\,\chi_{[-N,N]}\}$converges to f pointwise as $N \to \infty$, and each $f_N \in (L^1 \bigcap L^2)(\Re)$.

Lemma 2.7.6 [17]: Let $\{f_N = f\,\chi_{[-N,N]}\}$, then $\{f_N\}$ is a Cauchy sequence in the norm of $L^2(\Re)$ and $\lim\limits_{N\to\infty}\|f - f_N\|_{L^2} = 0$.

Proof : given any $\varepsilon > 0, \exists$ a step function s_m such that $\|f - s_m\|_2^2 < \varepsilon/2$, choose N so large that the support of s_m is contained in $[-N, N]$, then

$$\|s_m - f_N\|_2^2 = \int_{-N}^{N}|s_m - f_N|^2\,dt \le \int_{-\infty}^{\infty}|s_m - f|^2\,dt = \|s_m - f\|_2^2 \quad ,$$

so,

$$\begin{aligned}\|f - f_N\| &= \|(f - s_m) + (s_m - f_N)\| \\ &\le \|(f - s_m)\| + \|(s_m - f_N)\| \\ &\le 2\|f - s_m\| < \varepsilon .\end{aligned}$$

Note that if $\{s_n\}$ is a Cauchy sequence of step functions that converges to f, then $F(\{s_n\})$ is also Cauchy sequence, so we can defined $F(f)$ by $F(f)=\lim\limits_{n\to\infty} F(\{s_n\})$. Moreover, the definition of $\hat{f}(w)$ for $L^2(\Re)$ functions doesn't depend on the choice of such sequence in $(L^1\bigcap L^2)(\Re)$, so any other Cauchy sequence from $(L^1\bigcap L^2)(\Re)$ that approximate $f\in L^2(\Re)$ can be used to define $F(f)$ like $\{f_N\}$.

Theorem 2.7.7 [12]: If $f\in L^2(\Re)$, $\hat{f}(w)=\lim\limits_{N\to\infty}\frac{1}{\sqrt{2\pi}}\int_{-N}^{N} f(x)e^{-iwx}dx$,

where the convergence in the L^2- norm.

Proof : by lemma 2.7.6 $\|f-f_N\|_2\to 0$, as $N\to\infty$ where f_N is the truncated functions have a Fourier transform given by $\hat{f}_N(w)=\frac{1}{\sqrt{2\pi}}\int_{-N}^{N} f(x)e^{-iwx}dx$.

So,

$$\left\|\hat{f}-\hat{f}_N\right\|_2=\|F(f-f_N)\|_2=\|f-f_N\|_2,$$

hence,

$$\lim_{N\to\infty}\left\|\hat{f}-\hat{f}_N\right\|_2=0.$$ The proof is complete.

Lemma 2.7.8 [12]: If $f\in L^2(\Re)$ and $g=\bar{\hat{f}}$, then $f=\bar{\hat{g}}$.

Theorem 2.7.9 [12]: Inversion formula for L^2- functions

If $f\in L^2(\Re)$, then $f(x)=\lim\limits_{n\to\infty}\frac{1}{\sqrt{2\pi}}\int_{-n}^{n}\hat{f}(w)e^{iwx}dw$

Where the convergence is respect to the L^2- norm.

Proof : If $f\in L^2(\Re)$ and $g=\bar{\hat{f}}$, by lemma 2.7.8

$$f=\bar{\hat{g}}=\lim_{n\to\infty}\frac{1}{2\pi}\overline{\int_{-n}^{n} e^{-iwt}g(w)dw}$$

$$=\lim_{n\to\infty}\frac{1}{2\pi}\int_{-n}^{n} e^{iwt}\,\overline{g(w)}\,dw$$

$$= \lim_{n\to\infty} \frac{1}{2\pi} \int_{-n}^{n} e^{iwt}\, \hat{f}(w)\, dw\,.$$

Corollary 2.7.10 [12]: If $f \in (L^1 \bigcap L^2)(\Re)$, then $f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{iwx} dw\,.$

Holds almost everywhere in $\Re$.

It's easy to show that the Fourier transform is one to one map of $L^2(\Re)$ on to itself. This ensures that every square integrable function is the Fourier transform of a square integrable function.

Parseval's Relation

The energy carried by a signal $f(x)$ is: $\int_{-\infty}^{\infty} |f(x)|^2 dt = \int_{-\infty}^{\infty} f(x)\overline{f(x)}\, dx$

Where

$$\overline{f(x)} = \overline{\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(w) e^{iwx}\, dx} = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-iwx}\, \overline{\hat{f}(w)}\, dx \;\;,$$

So, we have that,

$$\int_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x)\,\overline{\hat{f}(w)}\, e^{-iwx}\; dw\, dx$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\hat{f}(w)} \left[\int_{-\infty}^{\infty} f(x) e^{-iwx}\, dx \right] dw$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\hat{f}(w)}\, \hat{f}(w)\, dw = \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(w)|^2\, dw \;\;.$$

This formula $\int_{-\infty}^{\infty} |f(x)|^2 = dx \frac{1}{2\pi} \int_{-\infty}^{\infty} |\hat{f}(w)|^2 dw$, is called Parseval's Relation.

The general Parseval's Relation is defined by:

$< f, g > = \frac{1}{2\pi} < \hat{f}, \hat{g} >$, where $f, g \in L^2(\Re)$.

Theorem 2.7.11 [17]: Convolution Theorem

If f and g in $L^1(\Re)$, and the convolution between f and g is defined by $(f * g)(x) = \int_{-\infty}^{\infty} f(x-u) g(u) du$, where *: is the convolution operator. Then

The Fourier transform of the convolution $(f*g)(x)$ is the product of the Fourier transform of these functions.

Remarks [1]:

a. We can see that the convolution of a bounded function with an integrable function and the convolution of two square integrable functions produce a continuous function.

b. The convolution in $L^1(\Re)$ tends to make functions smoother but less localize, for example if f and g in $L^1(\Re)$ with compact support equal to say, $[-a,a]$ and $[-b,b]$, then the support of $(f*g)(x)$ will be equal to $[-(a+b),(a+b)]$.

Poisson Summation Formula

In many applications it is necessary to form a periodic function from a nonperiodic function with finite energy for the purpose of analyzing.

Poisson's summation formula is useful in relating the time-domain information of such a function with its spectrum.

Theorem 2.7.12 [12]: If $f \in L^1(\Re)$, then the series $\sum_{-\infty}^{\infty} f(x+2n\pi)$ converges absolutely for almost all $x \in (0,2\pi)$, and its sum

$$F(x) \in L^1(0,2\pi) \quad \text{with} \quad F(x+2n\pi) = F(x), \quad x \in \Re.$$

And, if a_n denotes the Fourier coefficient of *F*, then

$$a_n = \frac{1}{2\pi}\int_0^{2\pi} F(x)e^{-inx}\,dx = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(x)e^{-inx}dx = \frac{1}{2\pi}\hat{f}(n).$$

Proof: we have

$$\sum_{n=-\infty}^{\infty}\int_0^{2\pi}|f(x+2n\pi)|\,dx = \lim_{N\to\infty}\sum_{n=-N}^{N}\int_0^{2\pi}|f(x+2n\pi)|\,dx$$

$$= \lim_{N\to\infty}\sum_{n=-N}^{N}\int_{2n\pi}^{2(n+1)\pi}|f(t)|\,dt$$

$$= \lim_{N\to\infty} \int_{-2\pi N}^{2\pi(N+1)} |f(t)|\, dt$$

$$= \int_{-\infty}^{\infty} |f(t)|\, dt < \infty .$$

It follows from lebesgue theorem on monotone convergence that

$$\int_0^{2\pi} \sum_{n=-\infty}^{\infty} |f(x+2n\pi)|\, dx = \sum_{n=-\infty}^{\infty} \int_0^{2\pi} |f(x+2n\pi)|\, dx < \infty$$

hence, the series $\sum_{-\infty}^{\infty} f(x+2n\pi)$ converges absolutely for almost all x, and

$$F(x) = \sum_{n=-\infty}^{\infty} f(x+2n\pi) \in L^1(0,2\pi) \text{ with } F(x+2n\pi) = F(x), \quad x \in \Re$$

so, we consider the Fourier series of F given by $F(x) = \sum_{m=-\infty}^{\infty} a_m e^{imx}$, where the coefficient a_m is

$$a_m = \frac{1}{2\pi} \int_0^{2\pi} F(x) e^{-imx}\, dx = \frac{1}{2\pi} \int_0^{2\pi} (\lim_{N\to\infty} F_N(x)) e^{-imx}\, dx$$

$$= \lim_{N\to\infty} \frac{1}{2\pi} \int_0^{2\pi} \sum_{n=-N}^{N} f(x+2n\pi) e^{-imx}\, dx$$

$$= \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} \int_0^{2\pi} f(x+2n\pi) e^{-imx}\, dx$$

$$= \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=-N}^{N} \int_{2n\pi}^{2\pi(n+1)} f(t) e^{-imt}\, dt$$

$$= \lim_{N\to\infty} \frac{1}{2\pi} \int_{-2N\pi}^{2\pi(N+1)} f(t) e^{-imt} dt$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t) e^{-imt}\, dt = \frac{1}{2\pi} \hat{f}(m) .$$

Hence if the Fourier series of $F(x)$ converges to $F(x)$, then for $x \in \Re$

$$F(x) = \sum_{n=-\infty}^{\infty} f(x+2n\pi) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \hat{f}(n) e^{inx}$$

Put $x = 0$, the last formula becomes $\sum_{n=-\infty}^{\infty} f(2n\pi) = \sum_{n=-\infty}^{\infty} \frac{1}{2\pi} \hat{f}(n)$, which is called Poisson summation formula.

Sampling Theorem

One of the fundamental results in Fourier analysis is the Shannon sampling theorem which asserts that a band limited function can be recovered from its samples on a regularly spaced set of points in $\Re$ .This result is basic in continuous-to- digital signal processing.

Definition 2.7.13 [12]: A function f is said to be frequency band limited if there exist a constant $\Omega > 0$, such that $\hat{f}(w) = 0$ for $|w| > \Omega$.

When Ω is the smallest frequency for which the preceding equation is true, the natural frequency $\gamma := \dfrac{\Omega}{2\pi}$ is called the **Nyquist** frequency, and $2\gamma := \dfrac{\Omega}{\pi}$ is the **Nyquist** rate.

Theorem 2.7.14 [4]: Shannon – Whittaker sampling theorem

Suppose that $\hat{f}(w)$ is piecewise smooth continuous, and that $\hat{f}(w) = 0$ for $|w| > \Omega$.

Then f is completely determined by its value at the point $t_j = \dfrac{j\pi}{\Omega}$, $j = 0, \pm 1, \pm 2, \ldots$

More precisely, f has the following series expansion

$$f(x) = \sum_{j=-\infty}^{\infty} f\left(\frac{j\pi}{\Omega}\right) \frac{\sin(\Omega x - j\pi)}{\Omega x - j\pi} \quad ,$$

where the series converge uniformly.

Proof: expand $\hat{f}(w)$ as a Fourier series on the interval $[-\Omega, \Omega]$

$$\hat{f}(w) = \sum_{k=-\infty}^{\infty} c_k \, e^{\frac{i\pi k w}{\Omega}} \quad , \quad c_k = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} \hat{f}(w) e^{\frac{-ik\pi w}{\Omega}} \, dw$$

since $\hat{f}(w) = 0$ for $|w| \geq \Omega$,then

$$c_k = \frac{\sqrt{2\pi}}{2\Omega} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(w) e^{\frac{-ik\pi w}{\Omega}} \, dw$$

By theorem 2.7.2, $c_k = \frac{\sqrt{2\pi}}{2\Omega} f\left(\frac{-k\pi}{\Omega}\right)$, so by changing the summation index from k to $j=-k$, and using the expression for c_k, we obtain

$$\hat{f}(w) = \sum_{j=-\infty}^{\infty} \frac{\sqrt{2\pi}}{2\Omega} f\left(\frac{j\pi}{\Omega}\right) e^{\frac{-i\pi jw}{\Omega}}.$$

Since $\hat{f}(w)$ is continuous, piecewise smooth function the last series is converge uniformly.

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \hat{f}(w) e^{iwt}\, dw, \text{ since } \hat{f}(w) = 0 \text{ for } |w| \ge \Omega$$

by some calculation we have

$$f(x) = \sum_{j=-\infty}^{\infty} \frac{\sqrt{2\pi}}{2\Omega} f\left(\frac{j\pi}{\Omega}\right) \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} e^{\frac{-i\pi jw}{\Omega}+iwx}\, dw \quad \text{but} \quad \int_{-\Omega}^{\Omega} e^{\frac{-i\pi jw}{\Omega}+iwx}\, dw = 2\Omega \frac{\sin(x\Omega - j\pi)}{(x\Omega - j\pi)}$$

So,

$$f(x) = \sum_{j=-\infty}^{\infty} f\left(\frac{j\pi}{\Omega}\right) \frac{\sin(\Omega x - j\pi)}{\Omega x - j\pi}.$$

The convergence rate in the last series is slow since the coefficient in absolute value decay like $\frac{1}{j}$. The convergence rate can be increased so that the terms behaves like $\frac{1}{j^2}$, by a technique called **Over sampling**.

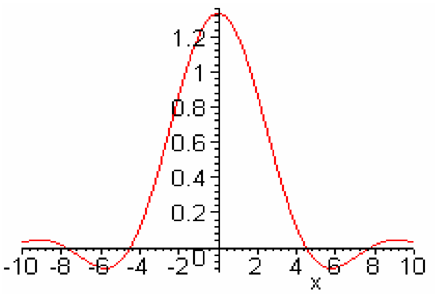
If a signal is sampled below the Nyquest rate, then the signal reconstructed will not only missing high frequency components transferred to low frequencies that may not have been in the signal at all. This phenomenon is called **aliasing.**

Example 2.7.4:

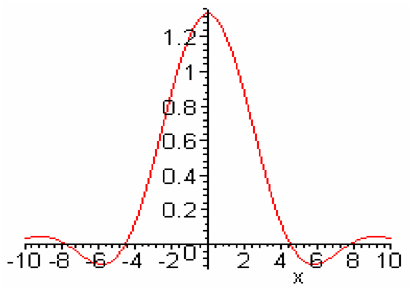
Consider the function f defined by $\hat{f}(w) = \begin{cases} \sqrt{2\pi}(1-w^2) & \text{if } |w| \le 1 \\ 0 & \text{if } |w| > 1 \end{cases}$

$f(x) = \frac{4\sin x - 4x\cos x}{x^3}$. The plot of f is given in Figure (13).

Since $\hat{f}(w)=0$ for $|w|>1$, the frequency Ω from the sampling theorem can be chosen to be any number that is greater than or equal to 1. With Ω =1, we graph the partial sum of the first 30 terms in the series given in the sampling theorem in Figure (13); note that the two graph are nearly identical.



(a): f



(b): S_{30}

Figure 13

Chapter three

Wavelets Analysis

3.1. Introduction

3.2. Continuous Wavelet Transform

3.3. Wavelet Series

3.4. Multiresolution Analysis (MRA)

3.5. Representation of functions by Wavelets

Chapter 3

Wavelets Analysis

3.1 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. They have advantages over traditional Fourier methods in analyzing physical situations where the signal contains discontinuities and sharp spikes. Like Fourier analysis, wavelet analysis deals with expansion of functions in terms of a set of basis functions. Unlike Fourier analysis, wavelet analysis expands functions not in terms of trigonometric polynomials but in terms of wavelets, which are generated in the form of translations and dilations of a fixed function cared the mother wavelet.

3.2 Continuous Wavelet Transform

The continuous wavelet transform (CWT) provides a method for displaying and analyzing characteristic of signals that are dependent on time and scale. The CWT is similar to the Fourier transform in the since that its based on a single function ψ and that this function is scaled. But unlike the Fourier transform, we also shift the function, thus, the CWT is an operator that takes a signal and produces a function of two variables: time and scale, as a function of two variables, it can be considered as surface or image.

In this section, we give formal definitions of wavelet and CWT of a function, and the basic properties of them. In addition, we will introduce the inversion formula for the CWT as in case for the Fourier transform. The CWT is defined with respect to a particular function, called mother wavelet, which satisfies some particular properties. As the kernel function of a signal transform, its important that the mother wavelet be designed so that the transform can be inverted. Even if the application of the CWT doesn't require such transform inversion, the invertibility of the CWT is necessary to ensure that no signal information is lost in the CWT.

Definition 3.2.1 [12]: Integral wavelets transform

If $\psi \in L^2(\Re)$ satisfies the admissibility condition $C_\psi := \int_{-\infty}^{\infty} \frac{|\hat{\psi}(w)|^2}{|w|} dw < \infty$, then ψ is called basic wavelet or mother wavelet.

Relative to every mother wavelet, the integral wavelet transform on $L^2(\Re)$ is defined by: $(W_\psi f)(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(x) \overline{\psi}\left(\frac{x-b}{a}\right) dx$, $f \in L^2(\Re)$.

Where $a, b \in \Re$.

The most important property that must be satisfied by mother wavelet is the admissibility condition which required for an inverse wavelet transform to exist. We suppose that ψ is continuous with continuous Fourier transform, if $\hat{\psi}(0) \neq 0$, then from continuity there is small interval I containing 0, and $\varepsilon > 0$ such that $|\hat{\psi}(w)| > \varepsilon,\ \forall w \in I$ but it would be followed

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(w)|^2}{|w|} dw \geq \int_I \frac{|\hat{\psi}(w)|^2}{|w|} dw \geq \int_I \frac{\varepsilon^2}{|w|} dw = \infty .$$

The admissibility condition therefore implies that $\hat{\psi}(0)=0$ or $\int_{-\infty}^{\infty}\psi(x)dx=0$,

for this to occur the mother wavelet must contain oscillations, it must have sufficient negative area to cancel out the positive area.

Example 3.2.1: Haar wavelet

The Haar wavelet is one of the classic example defined by

$$\psi(x)=\begin{cases}1 & ,\ 0\le x\le \frac{1}{2}\\ -1 & ,\ \frac{1}{2}\le x\le 1\\ 0 & ,\ \text{otherwise}\end{cases}$$

The Haar wavelet has compact support, and clearly $\int_{-\infty}^{\infty}\psi(x)dx=0$, and $\psi\in L^2(\Re)$,But this wavelet is not continuous, its Fourier transform is given by

$$\hat{\psi}(w)=ie^{\frac{-iw}{2}}\frac{\sin^2(w/4)}{(w/4)}$$

where

$$C_\psi := \int_{-\infty}^{\infty}\frac{|\hat{\psi}(w)|^2}{|w|}dw=16\int_{-\infty}^{\infty}|w|^{-3}\left|\sin\frac{w}{4}\right|^4 dw<\infty .$$

Both ψ and $\hat{\psi}$ are plotted in Figure 1, 2 respectively.

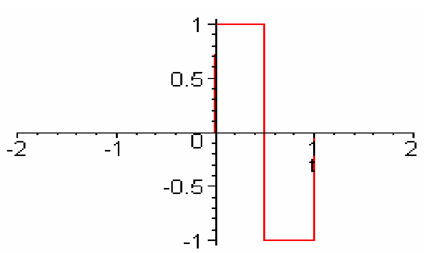


Figure 1

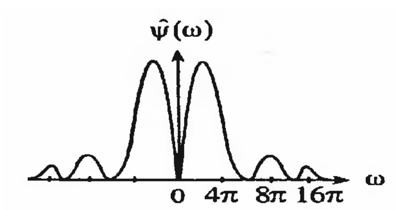


Figure 2

These Figures indicate that the Haar wavelet has good time localization but poor frequency localization, and this because the function $|\hat{\psi}(w)|$ is even and decays slowly as $\frac{1}{w}$ as $w \to \infty$, which means that it doesn't have compact support in the frequency domain.

Most of applications of wavelets exploit their ability to approximate functions as efficiently as possible, that is few coefficients as possible, so in addition to the admissibility condition, there are other properties that may be useful in particular application [1].

Localization property: we want ψ to be well localized in both time and frequency. In other word, ψ and its derivative must decay very rapidly. For frequency localization $\hat{\psi}(w)$ must decay sufficiently rapidly as $w \to \infty$, and $\hat{\psi}(w)$ should be flat in the neighborhood of $w = 0$. The flatness at $w = 0$ is associated with the number of vanishing moments of ψ. A wavelet is said to be M vanishing moment if $\int_{-\infty}^{\infty} \psi(x) x^m \, dx = 0$, $m = 0, 1, \quad , M-1$.

Wavelets with large number of vanishing moment result in more flatness when frequency w is small.

Smoothness: The smoothness of the wavelet increase with the number of vanishing moment.

Compact support: We say that ψ has compact support on I if its vanish outside these interval. If ψ has M vanishing moment, then its support is at least of length $2M$-1, so the Haar wavelet has minimum support equal to 1. Also, [The smoother wavelet, the longer support] this relation implies that there is no orthogonal wavelet that is C^{∞} and has compact support.

Example 3.2.2: consider the sinc wavelet system

$\varphi = \sin(nx)/nx$, where φ is the scaling function. The corresponding mother wavelet $\psi = 2\varphi(2x) - \varphi(x)$.

This wavelet has infinite number of vanishing moment and hence has infinite support see Figure 3.

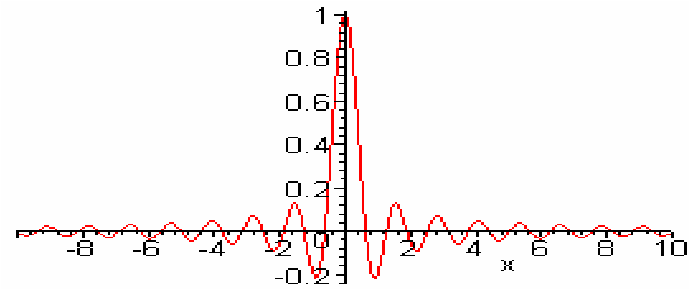


Figure 3

Theorem 3.2.2 [12]: If ψ is a wavelet and ϕ is bounded integrable function, then the convolution function $\psi * \phi$ is a wavelet.

Note that we can use theorem 3.2.2 to generate other wavelets, for example smooth wavelet.

Example 3.2.3: The convolution of the Haar wavelet with the function $\phi(x) = e^{-x^2}$, generate smooth wavelet, as shown in Figure 4.

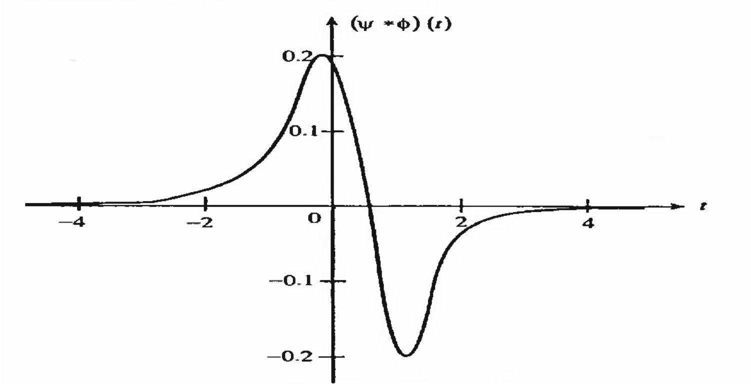


Figure 4

Example 3.2.4: Mexican hat wavelet

Its defined by the second derivative of a Gaussian function $\psi(x) = (1 - x^2)e^{\frac{-x^2}{2}}$, where $\hat{\psi}(w) = \sqrt{2\pi}\, w^2 e^{\frac{-w^2}{2}}$, see Figure 5, 6 related to ψ and $\hat{\psi}$ respectively.

This wavelet is smooth, and has two vanishing moment. In the contrast of the Haar wavelet, this wavelet has excellent localization in both time and frequency domain.

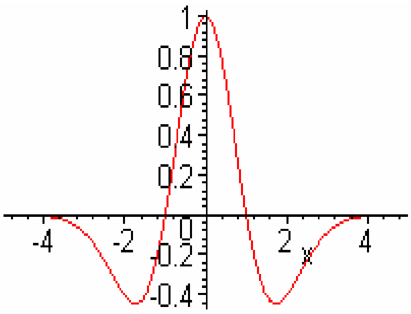


Figure 5

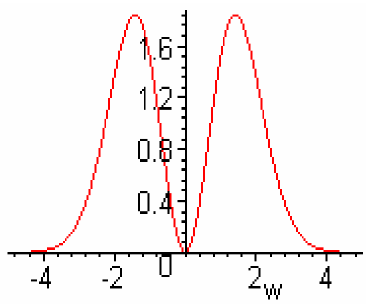


Figure 6

Basic property of wavelet transform

The following theorem gives several properties of CWT.

Theorem 3.2.3 [12]: If ψ and Φ are wavelets, and let $f, g \in L^2(\Re)$, then

1. Linearity , $W_\psi(\alpha f + \beta g) = \alpha(W_\psi f) + \beta(W_\psi g)$. $\alpha, \beta \in \Re$

2. Translation, $W_\psi(T_c f) = (W_\psi f)(a, b-c)$.
3. Dilation, $W_\psi(D_c f) = \frac{1}{\sqrt{c}}(W_\psi f)\left(\frac{a}{c}, \frac{b}{c}\right) \quad , c > 0$.
4. Symmetry, $W_\psi(f) = \overline{W_f \psi\left(\frac{1}{a}, \frac{-b}{a}\right)} \quad , a \neq 0$.
5. Antilinearity, $W_{(\alpha\psi + \beta\Phi)} f = \overline{\alpha}(W_\psi f) + \overline{\beta}(W_\psi g)$.

Theorem 3.2.4 [12]: Parsival's formula for wavelet transform

If $\psi \in L^2(\Re)$ and $W_\psi f(a,b)$ is the wavelet transform of f, then for any $f, g \in L^2(\Re)$

$$C_\psi < f, g > = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_\psi f(a,b) \overline{W_\psi g(a,b)} \frac{dbda}{a^2} \tag{3.2.1}$$

where

$$C_\psi := \int_{-\infty}^{\infty} \frac{|\hat{\psi}(w)|^2}{|w|} dw < \infty .$$

Proof: By Parsival's relation for the Fourier transforms, we have

$$(W_\psi f)(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(x) \overline{\psi}\left(\frac{x-b}{a}\right) dx$$

$$= < f, \psi_{a,b} >$$

$$= \frac{1}{2\pi} < \hat{f}, \hat{\psi}_{a,b} >$$

$$= \frac{1}{2\pi} \cdot \sqrt{|a|} \int_{-\infty}^{\infty} \hat{f}(w) e^{ibw} \overline{\hat{\psi}(aw)} dw \tag{3.2.2}$$

Similarly,

$$\overline{(W_\psi g)(a,b)} = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} \overline{g(x)} \psi\left(\frac{x-b}{a}\right) dx$$

$$= \frac{1}{2\pi} \cdot \sqrt{|a|} \int_{-\infty}^{\infty} \overline{\hat{g}(\sigma)} e^{-ib\sigma} \hat{\psi}(a\sigma) d\sigma . \tag{3.2.3}$$

Substituting (3.2.2) and (3.2.3) in the lift-hand side of (3.2.1) gives

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_\psi f(a,b) \overline{W_\psi g(a,b)} \frac{dbda}{a^2}$$

$$= \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{db\,da}{a^2}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}|a|\ \hat{f}(w)\,\overline{\hat{g}}(\sigma)\,\overline{\hat{\psi}}(aw)\,\hat{\psi}(a\sigma)\times\exp\{ib(w-\sigma)\}dw\,d\sigma\ .$$

Which is, by interchanging the order of integration,

$$= \frac{1}{(2\pi)}\int_{-\infty}^{\infty}\frac{da}{|a|}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(w)\,\overline{\hat{g}}(\sigma)\,\overline{\hat{\psi}}(aw)\,\hat{\psi}(a\sigma)\,dw\,d\sigma\times\frac{1}{2\pi}\int_{-\infty}^{\infty}\exp\{ib(w-\sigma)\}db$$

$$= \frac{1}{(2\pi)}\int_{-\infty}^{\infty}\frac{da}{|a|}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\hat{f}(w)\,\overline{\hat{g}}(\sigma)\,\overline{\hat{\psi}}(aw)\,\hat{\psi}(a\sigma)\,\delta(\sigma-w)\,dw\,d\sigma$$

$$= \frac{1}{(2\pi)}\int_{-\infty}^{\infty}\frac{da}{|a|}\int_{-\infty}^{\infty}\hat{f}(w)\,\overline{\hat{g}}(w)\,|\hat{\psi}(aw)|^2\,dw$$

which is, again interchanging the order of integration and putting $aw = x$,

$$= \frac{1}{(2\pi)}\int_{-\infty}^{\infty}\frac{|\hat{\psi}(x)|^2}{|x|}dx.\int_{-\infty}^{\infty}\hat{f}(w)\,\overline{\hat{g}}(w)\ dw\ .$$

$$= C_{\psi}\cdot\frac{1}{2\pi} < \hat{f}(w),\hat{g}(w) > .$$

Inversion formula

In chapter 2 we shown that the inversion formula for *f* can be written as $f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(w)e^{iwt}dw$, and this formula express the fact that *f* can be written as weighted sum of its various frequency component. The wavelet transform and its associated inversion formula also decompose a function in to weighted sum of its various frequency component. The difference between them that the wavelet inversion formula, two parameter *a* and *b* are involved since the wavelet transform involves *a* measure of frequency of *f* near the point $x = b$.

Theorem 3.2.5 [4]: Inversion formula

Suppose ψ is continuous wavelet satisfying the following

a. ψ has exponential decay, $\psi \in L^2(\Re)$.

b. $\int_{-\infty}^{\infty}\psi(x)dx = 0$.

Then for any $f \in L^2(\Re)$, f can be reconstruct by the formula

$$f(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |a|^{-1/2} W_\psi f(a,b) \psi\left(\frac{x-b}{a}\right) \frac{dbda}{a^2},$$

where the equality holds almost every where.

Proof: Let $G(x)$ be the quantity given on the right of the main statement of the theorem; that is,

$$G(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |a|^{-1/2} W_\psi f(a,b) \psi\left(\frac{x-b}{a}\right) \frac{dbda}{a^2} \qquad (3.2.4)$$

we must show that $G(x) = f(x)$.

By applying Plancherel's formula, which state that $\int uv = \int \overline{F(\overline{u})} F(v)$ to the b-integral occurring in the definition of $G(x)$ and where $v(b) = W_\psi f(a,b)$ and $u(b) = \psi\left(\frac{x-b}{a}\right)$, we can rewrite (3.2.4) as

$$G(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \frac{da}{a^2\sqrt{|a|}} \int_{-\infty}^{\infty} F_b\{W_\psi f(a,b)\}(y) \overline{F_b\left\{\overline{\psi\left(\frac{x-b}{a}\right)}\right\}(y)}\ dy \qquad (3.2.5)$$

where $F\{.\}$ stands for the Fourier transform of the quantity inside the brackets $\{.\}$, with respect to the variable b.

In order to apply the Plancherel's theorem, both of these functions must belong to $L^2(\Re)$. If f and ψ have finite support, then the b-support of $W_\psi f(a,b)$ will also be finite and so $W_\psi f(a,b)$ and $\psi\left(\frac{x-b}{a}\right)$ are L^2 functions in b. But

$$\overline{F_b\left\{\overline{\psi\left(\frac{x-b}{a}\right)}\right\}(y)} = a e^{iyx} \hat{\psi}(ay) \qquad (3.2.6)$$

$$F_b\{W_\psi f(a,b)\}(y) = a\sqrt{\frac{2\pi}{|a|}}\, \overline{\hat{\psi}(ay)}\, \hat{f}(y) \qquad (3.2.7)$$

Substitute (3.2.6) and (3.2.7) in (3.2.5), we obtain

$$G(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \frac{\sqrt{2\pi}\, da}{|a|} \int_{-\infty}^{\infty} |\hat{\psi}(ay)|^2\, \hat{f}(y) e^{iyx} dy$$

$$= \frac{1}{C_\psi} \int_{-\infty}^{\infty} \sqrt{2\pi}\, \hat{f}(y) e^{iyx} dy \int_{-\infty}^{\infty} \frac{|\hat{\psi}(ay)|^2}{|a|} da \qquad (3.2.8)$$

Where the last equality follows by interchanging the order of the *y*- and *a*-integrals.To calculate the *a*- integral on the right, we make a change of variables $u = ay$ provided that $y \neq 0$ to obtain

$$\int_{-\infty}^{\infty} \frac{|\hat{\psi}(ay)|^2}{|a|} da = \int_{-\infty}^{\infty} \frac{|\hat{\psi}(u)|^2}{|u|} du$$

$$= \frac{C_\psi}{2\pi}. \qquad (3.2.9)$$

Now, substitute (3.2.9) into (3.2.8) to obtain

$$G(x) = \frac{1}{C_\psi} \int_{-\infty}^{\infty} \sqrt{2\pi}\, \hat{f}(y) e^{iyx} \frac{C_\psi}{2\pi} dy$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(y) e^{iyx} dy = f(x).$$

where the last equality follows from the Fourier inversion theorem. This finish the proof.

3.3 Wavelet Series

It has been stated in section 3.2 that the continuous wavelet transform is a two-parameter representation of a function. In many applications, especially in signal processing, data are represented by a finite number of values, so it is important and often useful to consider discrete version of the continuous wavelet transform.

Basis for $L^2(\Re)$.

Note that any periodic function $f \in L^2([0,2\pi])$ can be expand as Fourier series: $f(x) = \sum_{n=-\infty}^{\infty} c_n e^{inx}$,where c_n is the Fourier coefficient of f , and we show that the equality hold if the system $\{e^{inx}\}_{n=0}^{\infty}$ is a complete orthonormal system. Now we consider to look for a basis for $L^2(\Re)$. Since every function in $L^2(\Re)$ must decay to zero at $\pm\infty$, the trigonometric function do not belong to $L^2(\Re)$. In fact, if we look for basis (waves) that generate $L^2(\Re)$, these waves should decay to zero at $\pm\infty$. Three simple operators on functions defined on $\Re$ play an important role in measure theory: translation, dilation, and modulation. We can apply some of these operators to construct orthonormal basis of $L^2(\Re)$ from single function in $L^2(\Re)$ say ψ. These basis are defined by $\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$, where the factor $2^{j/2}$ is to ensure the normalization of $\psi_{j,k}$ [6].

Definition 3.3.1 [3,12]: Orthonormal wavelet

A function $\psi \in L^2(\Re)$ is called an orthonormal wavelet, if the family $\{\psi_{j,k}\}$ is an orthonormal basis of $L^2(\Re)$.

There are several advantages to requiring that the scaling functions and wavelets be orthogonal. Orthogonal basis functions allow simple calculation of expansion coefficients and have Parseval's theorem that allows a partitioning of the signal energy in the wavelet transform domain.

Haar wavelets

The simplest example of an orthonormal wavelet is the classic Haar wavelet. It was introduced by Haar in 1910 in his PhD thesis. Haar's motivation was to find a basis of $L^2([0,1])$ that unlike the trigonometric system, will provide uniform convergence to the partial sums for continuous functions on [0,1]. This property is shared by most wavelets, in contrast with the Fourier basis for which the best we can expect for continuous functions is pointwise convergence *a.e*. There are two functions that play a primary role in wavelet analysis, the scaling function φ and the wavelet. These two functions generate a family of functions that can be used to break up or reconstruct a signal.

For the Haar system, let the scaling function be $\varphi = \begin{cases} 1 & ,0 \le x < 1 \\ 0 & ,\text{otherwise} \end{cases}$, see Figure 7

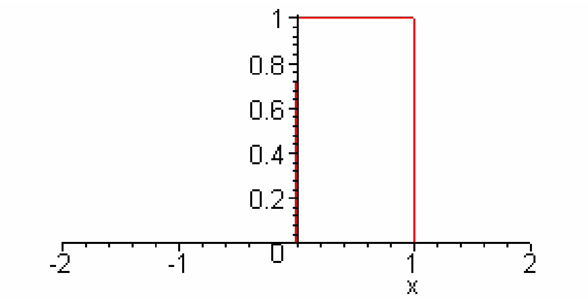


Figure 7

Let $V_0 = \text{span}\left(\{\varphi(x - k)\}_{k\in Z}\right)$ consists of all piecewise constant functions whose discontinuities are contained in the set of integers. Likewise, the subspaces $V_j = \text{span}\left(\{\varphi(2^j x - k)\}_{k\in Z}\right)$ are piecewise constant functions with jumps only at the integer multiples of 2^{-j}. Since *k* range over a finite set, each element of V_j is zero outside a bounded set. Such a function is said to have finite or compact support.

There are some basic properties of φ which are [4]:

a. $f(x) \in V_0$ iff $f\left(2^j x\right) \in V_j$ and $f(x) \in V_j$ iff $f\left(2^{-j} x\right) \in V_0$.

b. $\{\varphi(x-k)\}_{k\in Z}$ is an orthonormal basis for V_0, and $\{2^{j/2}\varphi(2^j x - k)\}_{k\in Z}$ is an orthonormal basis for V_j .

One way to construct ψ, by decompose V_j as an orthogonal sum of V_{j-1} and its complement. Start with $j=1$ and identify the orthogonal complement of V_0 in V_1 ,two key facts are needed to construct ψ [4]:

a. $\psi \in V_1$ and ψ can be express as $\sum_{k\in Z} a_k \varphi(2x-k)$ for some choice of a_k.

b. ψ is orthogonal to V_0, i.e. $\int_{-\infty}^{\infty} \psi(x)\varphi(x-k)dx = 0 \quad ,\forall k \in Z$.

The simplest ψ satisfying above condition is the function whose graph appears in Figure 1; this function can be written as $\psi(x) = \varphi(2x) - \varphi(2x-1)$ and is called the Haar wavelet.

Note that any function in V_1 is orthogonal to V_0 iff it is in $W_0 = \text{span}(\{\psi(x-k\}_{k\in Z})$.In otherworld, $V_1 = V_0 \oplus W_0$. In a similar manner, the following more general result can be established.

Theorem 3.3.2 [4]: Let W_j be the space of functions of the form

$$\sum_{k\in Z} a_k \psi(2^j x - k) \qquad a_k \in \Re$$

where we assume that only a finite number of a_k are zero. W_j is the orthogonal complement of V_j in V_{j+1} and $V_{j+1} = V_j \oplus W_j$.

Moreover, The wavelet $\{\psi_{j,k}\}$ form an orthonormal basis for W_j.

So, we can rewrite V_j as:

$$V_j = V_{j-1} \oplus W_{j-1} = V_{j-2} \oplus W_{j-2} \oplus W_{j-1}$$

$$= V_0 \oplus W_0 \oplus \qquad \oplus \mathrm{W}_{j-2} \oplus \mathrm{W}_{j-1}$$

and hence, the following theorem hold.

Theorem 3.3.3 [4]: The space $L^2(\Re)$ can be decomposed as an infinite orthogonal direct sum $\quad \mathrm{L}^2(\Re) = V_0 \oplus W_0 \oplus \qquad \oplus \mathrm{W}_{j-2} \oplus \mathrm{W}_{j-1}$

The most useful class of scaling functions are those that have compact support, the Haar scaling function is a good example of a compactly support function. The disadvantage of the Haar wavelets is that they are discontinuous and therefore do not approximate continuous functions very well. What is needed is a theory similar to what has been described above but with continuous versions of our building blocks, ψ and φ. The result theory, due to **Stephen Mallat** is called a multiresolution analysis.

3.4 Multiresolution Analysis (MRA)

The concept of multiresolution is intuitively related to the study of signals or images at different levels of resolution. The resolution of a signal is a qualitative description associated with its frequency content.

In 1986, Stephane Mallat and Yves Meyer first formulated the idea of multiresolution analysis in the context of wavelet analysis. This is a new and remarkable idea which deals with a general formalism for construction of an orthogonal bases of wavelets. Indeed, multiresolution analysis is central to all constructions of wavelets basis.

Mathematically, the fundamental idea of multiresolution analysis is to represent a function f as a limit of successive approximations, each of which is a finer version of the function f. These successive approximations correspond to different levels of resolutions. Thus, multiresolution analysis is a formal approach to constructing orthogonal wavelet bases using a definite set of rules and procedures.

Definition 3.4.1 [6]: Multiresolution Analysis

Let $\{V_j\}_{j\in Z}$ be a sequence of subspaces functions in $L^2(\Re)$ is called MRA with scaling function φ if the following conditions hold:-

1. (*Nested*), $V_j \subset V_{j+1}$ $\forall j \in$ Z.
2. (*Scaling*), $f \in V_j$ iff $f\left(2^{-j} x\right) \in V_0$ $\forall j \in$ Z.
3. (*Separation*), $\bigcap_{j\in Z} V_j = \{0\}$.
4. (*Density*), $\overline{\bigcup_{j\in Z} V_j} = L^2(\Re)$.
5. There exists a function $\varphi \in V_0$ such that $\{\varphi(x-k)\}_{k\in Z}$ is an orthonormal basis for V_0, that is,

$$\|f\|^2 = \int_{-\infty}^{\infty} |f|^2 dx = \sum_{k=-\infty}^{\infty} \left|< f, \varphi_{0,k} >\right|^2 \quad \forall f \in V_0 .$$

Sometimes, condition 5 is relaxed by assuming that $\{\varphi(x-k)\}_{k\in Z}$ is Riesz basis for V_0, that is for every $f \in V_0$, there exists a unique sequence $\{C_k\}_{k\in Z} \in l^2(Z)$ such that $f(x) = \sum_{k\in Z} C_k \varphi(x-k)$, with convergence in $L^2(\Re)$, and there exist two positive constant A and B independent of $f \in V_0$ such that

$$A\sum_{k\in Z} |C_k|^2 \le \|f\|^2 \le B\sum_{k\in Z} |C_k|^2$$

where $0 < A < B < \infty$.

In this case, we have a MRA with Riesz basis and we can then use φ to obtain a new scaling function $\tilde{\varphi}$ for which $\{\tilde{\varphi}(x-k)\}_{k\in Z}$ is orthonormal.

Example 3.4.1: The collection of subspaces V_j, consisting of the space of piecewise constant functions of finite support whose discontinuities are contained in the set of integer multiple of 2^{-j}, together with the Haar scaling function φ, satisfies the definition of MRA.

Example 3.4.2: Shannon multiresolution analysis

Here V_j is the space of band-limited signals $f \in L^2(\Re)$, with frequency band contained in the interval $\left[-2^j\pi, 2^j\pi\right]$. The scaling function defined by

$$\varphi(x) = \begin{cases} 1 & x = 0 \\ \dfrac{\sin \pi x}{\pi x} & x \neq 0 \end{cases}$$

The Fourier transform of φ is given by $\hat{\varphi}(w) = \chi_{[-\pi,\pi]}(w)$.

Clearly, the Shannon scaling function doesn't have finite support. However, its Fourier transform has a finite support in the frequency domain and has good frequency localization.

We turn to a discussion of properties common to every multiresolution analysis; our first result is that $\{\varphi_{j,k}\}_{k\in Z}$ is an orthonormal basis for V_j.

Theorem 3.4.2 [4]: Suppose $\{V_j\}_{J\in Z}$ is a multiresolution analysis with scaling function φ. Then for any $j \in Z$, the set of functions

$$\left\{\varphi_{j,k}(x) = 2^{j/2}\varphi\left(2^j x - k\right)\right\}_{k\in Z}$$

is an orthonormal basis for V_j

Proof: [4].

We are now ready to state the central equation in MRA, the scaling relation, which is also called two-scale relation, since its relates $\varphi(x)$ and the translates of $\varphi(2x)$.

Theorem 3.4.3 [4]: scaling relation

Suppose $\{V_j\}_{J\in Z}$ is a multiresolution analysis with scaling function φ. Then the following scaling relation holds:

$$\varphi(x) = \sum_{k\in Z} p_k \varphi(2x - k) \qquad \text{where} \quad p_k = 2\int_{-\infty}^{\infty} \varphi(x)\overline{\varphi(2x-k)}\,dx$$

Moreover, we also have

$$\varphi\left(2^{j-1}x-l\right)=\sum_{k\in Z} p_{k-2l}\varphi\left(2^{j}x-k\right) .$$

proof: [4].

Example 3.4.3: The values of the p_k for the Haar system are

$$p_0 = p_1 = 1.$$

and all other p_k are zero.

Construction of wavelet from a multiresolution analysis

We now pass to the construction of orthonormal wavelets from an MRA. Let W_0 be the orthogonal complement of V_0 in V_1; that is, $V_1 = V_0 \oplus W_0$. If we dilate the elements of W_0 by 2^j, we obtain closed subspace W_j of V_{j+1}, such that $V_{j+1} = V_j \oplus W_j, \forall j \in Z$

Since

$$V_j \to \{0\} \text{ as } j \to -\infty \quad \Rightarrow \quad V_{j+1} = \bigoplus_{L=-\infty}^{j} W_L \quad \forall j$$

and

$$V_j \to L^2(\Re) \text{ as } j \to \infty \quad ,$$

we have $L^2(\Re) = \bigoplus_{L=-\infty}^{\infty} W_L$

To find an orthonormal wavelet, as in case of the Haar system, all we need to do is to find a function $\psi \in W_0$ such that $\psi(x-k)$ is an orthonormal basis for W_0. In fact, if this is the case, then $\left\{2^{j/2}\psi\left(2^j x-k\right)\right\}_{k\in Z}$ is an orthonormal basis for $W_j \quad \forall j \in Z$, and hence $\{\psi_{j,k}\}_{k,j\in Z}$ is an orthonormal basis for $L^2(\Re)$, which shows that ψ is an orthonormal basis for $\Re$.

The scaling relation can be used to construct the associated function ψ that generates W_j.

Theorem 3.4.4 [4]: Suppose $\{V_j\}_{j\in Z}$ is a multiresolution analysis with scaling function

$$\varphi(x)=\sum_{k\in Z} p_k\varphi(2x-k)$$

(p_k are the coefficients in theorem 3.4.3). Let W_j be the span of $\{\psi(2^j x-k)\}_{k\in Z}$, where

$$\psi(x)=\sum_{k\in Z}(-1)^k\overline{p_{1-k}}\varphi(2x-k) \quad \text{and} \quad \psi_{j,l}(x)=2^{-1/2}\sum_{k\in Z}(-1)^k\overline{p_{1-k+2l}}\varphi_{j+1,k}$$

Then $W_j\subset V_{j+1}$ is the orthogonal complement of V_j in V_{j+1}. Furthermore, $\{\psi_{j,k}(x):=2^{j/2}\psi(2^j x-k)\}_{k\in Z}$ is an orthonormal basis for W_j, and hence the set of all wavelets, $\{\psi_{j,k}\}_{k,j\in Z}$ is an orthonormal basis for $L^2(\Re)$.

Proof: [12].

Daubechies wavelet

The wavelet that we looked at so far, Haar, Shannon wavelets have all major drawbacks. Haar wavelets have compact support but are discontinuous. Shannon wavelets are very smooth but extend throughout the whole real line. These wavelets, together with a few others having similar properties, were the only ones available before Ingrid Daubechies discovered the hierarchy of wavelets that the Haar wavelet, which is the only discontinuous one. The other wavelets in the hierarchy are compactly supported and continuous. Wavelet with compact support have many interesting properties. They can be constructed to have a given number of derivatives and to have a given number of vanishing moments [4].

Example 3.4.4: The associated value of the p_k can be computed to be

$$p_0=\frac{1+\sqrt{3}}{4},\ p_1=\frac{3+\sqrt{3}}{4},\ p_3=\frac{3-\sqrt{3}}{4},\ p_4=\frac{1-\sqrt{3}}{4}\ .$$

Consequently, the Daubechies scaling function (see Figure 8) becomes

$$\varphi(x)=\frac{1+\sqrt{3}}{4}\varphi(2x)+\frac{3+\sqrt{3}}{4}\varphi(2x-1)+\frac{3-\sqrt{3}}{4}\varphi(2x-2)+\frac{1-\sqrt{3}}{4}\varphi(2x-3).$$

And the corresponding mother wavelet ψ is

$$\psi(x) = -\frac{1+\sqrt{3}}{4}\varphi(2x-1) + \frac{3+\sqrt{3}}{4}\varphi(2x) - \frac{3-\sqrt{3}}{4}\varphi(2x+1) + \frac{1-\sqrt{3}}{4}\varphi(2x+2).$$

and this is called the Daubechies wavelet, see Figure 9.

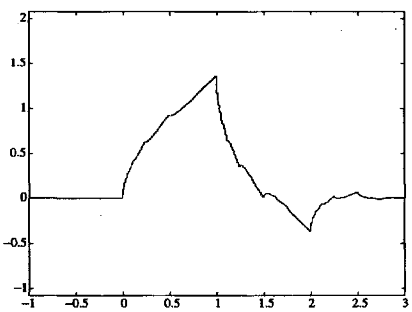


Figure: 8

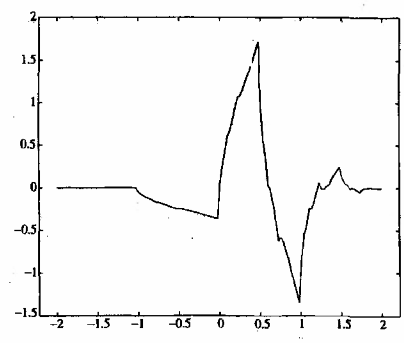


Figure: 9

3.5 Representation of functions by Wavelets

Since a wavelet system $\{\psi_{j,k}\}_{k,j\in Z}$ is an orthonormal basis for $L^2(\Re)$, we know that for any $f \in L^2(\Re)$,

$$f = \sum_{j\in Z}\sum_{k\in Z} < f, \psi_{j,k} > \psi_{j,k}$$

with convergence in the $L^2(\Re)$-norm.

The goal of most expansions of a function or signal is to have the coefficients of the expansion $a_{j,k}$ give more information about the signal than is directly obvious from the signal itself. A second goal is to have most of the coefficients be zero or very small. This is what is called a spare representation and is extremely important in applications for statistical estimation and data compression.

Although this expansion is called the discrete wavelet transform (DWT), it probably should be called a wavelet series since it is a series expansion which maps a function of a continuous variable into sequence of coefficients much the same way the Fourier series does.

This wavelet series expansion is in terms of two indices, the time translation *k* and the scaling index *j*. For the Fourier series, there are only two possible values of *k*, zero and $\pi/2$, which give the sine terms and the cosine terms. The values *j* give the frequency harmonics. In other words, the Fourier series is also a two-dimensional expansion, but that is not seen in the exponential form and generally not noticed in the trigonometric form.

The coefficients in this wavelet expansion are called the discrete wavelet transform of the signal *f*, these wavelet coefficients can be completely describe the original signal and can be used in a way similar to Fourier series coefficients for analysis, description, approximation, and filtering. If the wavelet system is orthogonal, these coefficients can be calculated by inner products.

The DWT is similar to a Fourier series but, in many ways, is much more flexible and informative. It can be made periodic like a Fourier series to represent periodic signals efficiently. However, unlike a Fourier series, it can be used directly on non-periodic transient signals with excellent results.

The main purpose of this section is to study if such expansions are well defined and converge in then setting of other function spaces. In particular we shall study the convergence in L^p-norm and the uniform convergence of wavelet expansions on the real line.

Convergence of the Haar series

We know that the Haar wavelet form an orthonormal basis for $L^2(\Re)$, then for any $f \in L^2(\Re)$, we have

$$f = \sum_{j,k \in Z} c_{j,k} \psi_{j,k} \tag{3.5.1}$$

is the Haar series of f, the Haar coefficients defined by

$$c_{j,k} =< f, \psi_{j,k} >= \int_{-\infty}^{\infty} f(x) \psi_{j,k}(x) dx$$

The completeness of $L^2(\Re)$ further assures that the series above converges in $L^2(\Re)$. In order to identify the sum of the Haar series, let $p_n, n \in Z$ be the projection operator of $f \in L^2(\Re)$ on to the space V_j defined by [14]:

$$p_n(f) = 2^n \int_{I_{kn}} f(y) dy \qquad \text{where } I_{kn} := \left((k-1)/2^n, k/2^n\right]$$

This formula can be written explicitly in terms of the Haar scaling function

$$p_n(f) = \int_{\Re} K_n(x,y) f(y) dy \quad ,$$

where

$$K_n(x,y) = 2^n \sum_{k \in Z} \varphi\left(2^n x - k\right) \varphi\left(2^n y - k\right) = \begin{cases} 2^n & , x, y \in I_{kn} \\ 0 & , \text{otherwise} \end{cases}$$

$\Rightarrow p_n$ increasing and converge to the identity function in the sense that

a. $p_n f = f$ implies $p_{n+1} f = f$.

b. $\lim_{n \to \infty} p_n f = f$ in $L^2(\Re)$.

Moreover, let $L_n(x,y) = K_{n+1}(x,y) - K_n(x,y)$, in terms of the Haar function we get

$$L_n(x,y) = 2^n \sum_{k \in Z} \psi\left(2^n x - k\right) \psi\left(2^n y - k\right)$$

So, we have

$$p_{n+1} f - p_n f = \sum_{k \in Z} \psi_{n,k}(x) \left(\int_{\Re} f(y) \psi_{n,k}(y) dy \right) \tag{3.5.2}$$

hence we can write the original projection operator in the form

$$p_{n+1} = p_0 f + \sum_{j=0}^{n} p_{j+1} f - p_j f$$

as $n \to \infty$, this yield to the one-sided Haar series representation

$$f = p_0 f + \sum_{j=0}^{\infty} \sum_{k \in Z} \psi_{j,k} \left(\int_{\Re} f(y) \psi_{j,k}(y) dy \right)$$

Completeness the Haar system

To prove the validity of the two-sided Haar series (3.5.1), we go back to (3.5.2) and write

$$p_{n+1} f - p_{-m} f = \sum_{j=-m}^{n} \sum_{k \in Z} c_{j,k} \psi_{j,k} \tag{3.5.3}$$

so, it remains to prove that $p_{-m} f \to 0$ and $p_{n+1} f \to f$ when $m, n \to \infty$.

First, we prove that the operators p_n have uniformly bounded operators norm.

Lemma 3.5.1 [14]: For any $f \in L^p(\Re), 1 \le p < \infty$, then $\|p_n f\|_p \le \|f\|_p, \forall n$.

Proof : For $p = 2$

From the definition of $p_n f$, we apply Cauchy-Schwarz inequality to obtain

$$|p_n f(x)|^2 \le 2^n \int_{I_{kn}} |f(y)|^2 dy \quad \text{for} \quad x \in I_{kn}$$

$$\Rightarrow \quad \int_{I_{kn}} |p_n f(x)|^2 dx \le \int_{I_{kn}} |f(x)|^2 dx$$

$$\Rightarrow \quad \int_{\Re} |p_n f(x)|^2 dx \le \int_{\Re} |f(x)|^2 dx .$$

For $p \ne 2$

Set $p' = \dfrac{p}{1-p}$, where $p > 1$, then Holder's inequality gives

$$|p_n f(x)| \le 2^n \left(\int_{I_{kn}} |f(y)|^p dy \right)^{\frac{1}{p}} 2^{-n/p'} \qquad , x \in I_{kn}$$

$$\Rightarrow \quad |p_n f(x)|^p \le 2^{np} \left(\int_{I_{kn}} |f(y)|^p dy \right) 2^{-np/p'}$$

$$\Rightarrow \quad \int_{I_{kn}} |p_n f(x)|^p dx \le 2^{-n} 2^{np} 2^{-np/p'} \int_{I_{kn}} |f(y)|^p dy = \int_{I_{kn}} |f(y)|^p dy$$

$$\Rightarrow \quad \int_{I_{kn}} |p_n f(x)|^p dx \le \int_{\Re} |f(x)|^p dx .$$

This proof also applies in case p = 1, by setting $\frac{1}{\infty} = 0$ whenever p' appears.

Let we define $C_0(\Re)$ to be the set of continuous functions vanishing at infinity and $C_{00}(\Re)$to be the set of continuous functions with compact support.

Lemma 3.5.2 [14]:

1. If $f \in C_0(\Re)$, we have $\|p_{-m} f\|_\infty \to 0$ as $m \to \infty$.
2. If $f \in L^2(\Re)$, we have $\|p_{-m} f\|_2 \to 0$ as $m \to \infty$.

Proof : 1. if $g \in C_{00}(\Re)$ has support in [-k, k], then

$$0 \le x \le 2^m \Rightarrow \quad |p_{-m} g(x)| = 2^{-m} \int_0^k |g| dx \to 0$$

and

$$-2^m \le x \le 0 \Rightarrow \quad |p_{-m} g(x)| = 2^{-m} \int_{-k}^{o} |g| dx \to 0 .$$

hence, $\|p_{-m} g\|_\infty \to 0$. But these functions are dense in $C_0(\Re)$; given $f \in C_0(\Re)$ and $\varepsilon > 0$, there exist $g \in C_{00}(\Re)$, $h \in C_0(\Re)$ such that f = g + h, with $\|h\|_\infty < \varepsilon$.

Then

$$\limsup_{m\to\infty} \|p_{-m} f\|_\infty \le \limsup_{m\to\infty} \|p_{-m} h\|_\infty < \varepsilon$$

since ε is arbitrary, this proves the required convergence.

2. If $f \in L^2(\Re)$, for any $\varepsilon > 0$, f = g + h, where g is continuous function with compact support in [-k, k], k > 0, and $\|h\|_2 < \varepsilon$. Then for $2^m > k$, we have

$$-2^m \le x \le 2^m \Rightarrow \ |p_{-m} g(x)| = 2^{-m} \int_{-k}^{k} |g| \, dx \le 2^{-m} \sqrt{2k} \|g\|_2 .$$

$$\|p_{-m} g(x)\|_2 \le \sqrt{4k}\, 2^{-m/2} \|g\|_2$$

$$\|p_{-m}f(x)\|_2 \le \|p_{-m}g\|_2 + \|p_{-m}h\|_2 \le \|p_{-m}g\|_2 + \varepsilon$$

$\Rightarrow \quad \limsup_{m\to\infty} \|p_{-m}f\|_2 \le \varepsilon$, since ε is arbitrary, we conclude that

$$p_{-m}f \to 0 \text{ as } m \to \infty \quad .$$

To prove that $p_n f \to f$ as $n \to \infty$, we first prove that this holds on the dense subset of $C_{00}(\Re)$.

Lemma 3.5.3 [14]: If $f \in C_{00}(\Re)$, then $p_n f \to f$ uniformly and in $L^2(\Re)$, when $n \to \infty$.

Proof : let f be supported in [-k, k], $k \ge 1$. Given $\varepsilon > 0$, from the uniform continuity of f, $\exists \delta > 0$ such that $|f(y) - f(x)| < \frac{\varepsilon}{k}$, whenever $|x - y| < \delta$. If $2^n < \delta$, we have

$$|p_n f(x) - f(x)| \le \frac{\varepsilon}{\sqrt{2k}} \le \varepsilon, \forall x.$$

Which prove the uniform convergence.

To prove the L^2-convergence

$$\int_{\Re} |p_n f(x) - f(x)|^2 dx \le \int_{-k}^{k} \frac{\varepsilon^2}{2k} dx \le \varepsilon^2 .$$

$$\Rightarrow \|p_n f - f\|_2 \to 0 \text{ as } n \to \infty.$$

So, by lemma 3.5.2 and 3.5.3, we have thus proved the L^2-convergence of the Haar series.

Haar series in C_0 and L^p spaces

We have treated the L^2-convergence of the Haar series, now; we discuss the uniform convergence in spaces of continuous functions, as well as the norm convergence in $L^p(\Re)$ $1 \le p < \infty$.

First, we treat the convergence in the space $C_0(\Re)$, by lemma 3.14, 3.15, we have

$p_{-m} f \to 0$ as $m \to \infty$ and $p_n f \to f$ as $n \to \infty$, respectively.

Since the space $C_0(\Re)$ contains as a dense subspace the set of $C_{00}(\Re)$.It remains to prove that the operators are uniformly bounded

Lemma 3.5.4 [14]: For any $f \in B_c(\Re)$,(the space of bounded continuous functions), we have $\|p_n f\| \le \|f\|_\infty$.

Proof: for $x \in I_{kn}$, we have $|p_n f(x)| \le 2^n \int\limits_{I_{kn}} |f(y)|\, dy \le \|f\|_\infty$

This leads to the following general proposition on uniform convergence.

Proposition 3.5.5 [14]: If $f \in C_0(\Re)$, then the Haar series (3.5.3) converge uniformly on the entire real line.

To prove the L^p-convergence, we must check that $p_{-m} f \to 0$ as $m \to \infty$ and $p_n f \to f$ as $n \to \infty$.

Lemma 3.5.6 [14]: Let $1 \le p < \infty$. Then $\|p_n f - f\|_p \to 0$ as $n \to \infty$.

Proof : The space $C_{00}(\Re)$ is dense in $L^p(\Re)$, from lemma 3.5.3 we have uniform convergence on this space. In particular if $\text{supp}(f) \subset [-k, k]$, for $n > N(\varepsilon)$, we have

$$\int\limits_{\Re} |p_n f(x) - f(x)|^P dx \le 2K\varepsilon^P .$$

which shows that $\|p_n f - f\|_p < \varepsilon (2k)^{\frac{1}{p}}$.

Lemma 3.5.7 [14]: Let $1 < p < \infty$, then $\|p_{-m} f\|_p \to 0$ when $m \to \infty$.

Proof : it suffices to check this for *g* continuous with compact support in [-*k*, *k*], if $2^m > k$, then

$$0 \le x \le 2^m \Rightarrow \quad |p_{-m} g(x)| = 2^{-m} \left| \int\limits_0^k g(y) dy \right|$$

$$\int\limits_0^\infty |p_{-m} g(x)|^p dx = 2^{-mp} 2^m \left| \int\limits_{-k}^k |g(y)|^p dy \right| .(2k)^{p/p'}$$

which tend to zero when $m \to \infty$. For $-2^m \le x \le 0$, we use the same fashion. Hence we can conclude the following.

Proposition 3.5.8 [14]: Let $1 < p < \infty$, for any $f \in L^p(\Re)$, the Haar series (3.5.3) converges in the norm of $L^p(\Re)$. And for $1 \le p < \infty$, the one-sided Haar series are hold.

Convergence of the wavelet expansion in $L^p(\Re)$

All the wavelet we will use in this subsection are assumed to arise from a multiresolution analysis (MRA). For the MRA we shall assume that the scaling function φ and the wavelet ψ have controlled decrease at infinity. Moreprecisely, there is a bounded function. $W:[0,\infty) \to \Re^+$, which is a radial decreasing L^1 – majorant of φ or ψ, if $|\varphi(x)| \le W(|x|)$ and W satisfies the following conditions [6]:-

1. $W \in L^1([0,\infty))$.
2. W is decreasing.
3. $\mathrm{W}(0) < \infty$.

Example 3.5.1: Two particularly natural choices for W are

$$W(|x|) = ce^{-\varepsilon|x|} \quad \text{for some } \varepsilon > 0$$

and

$$W(|x|) = \frac{c}{(1+|x|)^{\alpha}} \quad \text{for some } \alpha > 0.$$

Both examples are good majorants for the compactly support wavelets [6].

Suppose that we have a wavelet ψ arise from MRA with scaling function φ. Associated with the increasing sequence of subspaces $\{V_j\}_{j\in Z}$, we have the orthogonal projections of $L^2(\Re)$ onto V_j given by

$$p_j f = \sum_{k\in Z} < f, \varphi_{j,k} > \varphi_{j,k} \ , \ f \in L^2(\Re)$$

As in the Haar series representation we can write $< f, \varphi_{j,k} >$ as an integral and interchanging the order of summation and integration, to obtain

$$p_j f(x) = \int_{-\infty}^{\infty} 2^j K_\varphi\left(2^j x, 2^j y\right) f(y)dy \qquad (3.5.4)$$

where

$$K_\varphi(x,y) = \sum_{k\in Z} \varphi(x-k)\overline{\varphi(y-k)}$$

is the wavelet kernel.

Proposition 3.5.9 [6, 14]: The wavelet kernel $K_\varphi(x,y)$ enjoys the following properties:

1. $K_\varphi \in L^1_{loc}\left(\Re^2\right)$
2. $K_\varphi(x,y) = \overline{K_\varphi(x,y)}$
3. $\int_{\Re} \left|K_\varphi(x,y)\right| dy \le c < \infty$ and $\int_{\Re} K_\varphi(x,y)dy = 1$
4. $\left|K_\varphi(x,y)\right| \le cW\left(\frac{|x-y|}{2}\right).$ (3.5.5)

The main purpose is to prove that

a. $p_j f \to 0$ as $j \to -\infty$.

b. $p_j f \to f$ as $j \to \infty$.

Proposition 3.5.10 [6]: Suppose φ has radial decreasing L^1 –majorant W; then there exist $C > 0$ independent of j such that $\forall f \in L^p(\Re), 1 \le p \le \infty$**,** we have

$$\left\|p_j f\right\|_p \le C\|W\|_{L^1[0,\infty)}\|f\|_p$$

Proof : if $p = \infty$, by using (3.5.4) and (3.5.5) we get

$$\left|p_j f(x)\right| \le C\int_{\Re} 2^j W\left(\frac{2^j|x-y|}{2}\right)\left|f(y)\right|dy \le C\|f\|_\infty \|W\|_{L^1[0,\infty)}$$

If $p = 1$

$$\int_{\Re}\left|p_j f(x)\right|dx \le C\int_{\Re}\left(\int_{\Re}2^j W\left(\frac{2^j|x-y|}{2}\right)\left|f(y)\right|dy\right)dx$$

$$= C\int_{\Re}\left|f(y)\right|\left(\int_{\Re}2^j W\left(2^{j-1}|x-y|\right)dx\right)dy$$

$$\le C\|W\|_{L^1[0,\infty)}\,\|f\|_1.$$

These bounds allow us to formulate and prove a general theorem on the convergence of the scale projection operator.

Theorem 3.5.11 [14]: Suppose φ is the scaling function of an MRA which has radial decreasing L^1-majorant W

1- If $f \in B_{uc}(\Re)$, then $\|p_j f - f\|_\infty \to 0$ as $j \to \infty$.

2- If $f \in L^p(\Re)$, then $\|p_j f - f\|_p \to 0$ as $j \to \infty$, $1 \le p < \infty$

Proof : first we note that $p_j 1 = 1$, which follows from $\int_{\Re} K_\varphi(x,y)dy = 1$, this allows one to write $f - p_j f = 2^j \int_{\Re} K_\varphi\left(2^j x, 2^j y\right)\left[f(x) - f(y)\right]dy$, since f is continuous, given $\varepsilon > 0$, let $\delta > 0$ such that $\left|f(x) - f(y)\right| < \frac{\varepsilon}{2c}$ whenever $|x - y| < \delta$, so

$$f(x) - p_j f(x) = 2^j\left(\int_{|y-x|\le\delta} + \int_{|y-x|>\delta}\right)\left[f(x) - f(y)\right]K_\varphi\left(2^j x, 2^j y\right)dy$$

by apply the bound $\int_{\Re}\left|K_\varphi(x,y)\right|dy \le c$ in the first integral we conclude that this term is less than $\frac{\varepsilon}{2}, \forall j$. To estimate the second integral, we use the boundedness to obtain the upper bound

$$2\|f\|_\infty \int_{|y-x|>\delta} 2^j \left|K_\varphi\left(2^j x, 2^j y\right)\right| dy \le C\|f\|_\infty \int_{|y-x|>\delta} 2^j W\left(\frac{2^j|y-x|}{2}\right) dy$$

$$= C\|f\|_\infty \int_{\left|u-2^{j-1}x\right|>2^{j-1}\delta} W\left(\left|u-2^{j-1}x\right|\right) du$$

$$= C\|f\|_\infty \int_{|v|>2^{j-1}\delta} W(v)\, dv$$

which tend to zero as $j \to \infty$, by the dominate convergence theorem. This is a uniform bound independent of $x \in \Re$, from which we obtain the asserted uniform convergence.

To prove L^p convergence, we first discuss the case $p = 1$. From the uniform boundedness $\|p_j f\|_1 \le \|f\|_1$, it suffices to prove the L^p convergence on the dense set of continuous functions with compact support in [-R , R], for such f , we have

$$\|f - p_j f\|_1 = \int_{|x|\le 2R} |f(x)|\, dx + \int_{|x|>2R} |p_j f(x)|\, dx$$

The first integral tends to zero by virtue of the uniform convergence already proved, to estimate the second integral, we write

$$\int_{|x|>2R} |p_j f(x)|\, dx \le \|f\|_\infty \int_{|x|>2R} \left(\int_{|y|<R} 2^j K_\varphi\left(2^j x, 2^j y\right) dy \right) dx$$

$$\le C \int_{|y|<R} \int_{|x|>2R} 2^j W\left(\frac{2^j|x-y|}{2}\right) dx dy$$

$$\le C \int_{|y|<R} \int_{|x-y|>R} 2^j W\left(\frac{2^j|x-y|}{2}\right) dx dy$$

$$\le C \int_{|y|<R} \int_{|u|>R} 2^j W\left(\frac{2^j|u|}{2}\right) dx du$$

$$\le 2CR \int_{|v|>2^j R} W(v)\, dv \to 0 \text{ when } j \to \infty.$$

which complete the proof of L^p convergence.

To treat the case $1 < p < \infty$, it again suffices to deal with continuous functions with compact support. In this case we have the bounds

$$\left| f(x) - p_j f(x) \right|^p \leq \left\| f - p_j f \right\|_\infty^{p-1} \left| f(x) - p_j f(x) \right|$$

$$\int_{\Re} \left| f(x) - p_j f(x) \right|^p dx \leq \left\| f - p_j f \right\|_\infty^{p-1} \int_{\Re} \left| f(x) - p_j f(x) \right| dx$$

$$= \left\| f - p_j f \right\|_\infty^{p-1} \left\| f - p_j f \right\|_1 .$$

Which tends to zero, by the convergence in case $p = 1$, already proved. This proves the theorem.

Large scale analysis

To complete the analysis of L^p convergence of general wavelet series, it remains to prove that $p_j f \to 0$ as $j \to -\infty$. As in the case of Haar series, we expect only that this will take place for $L^p(\Re), 1 < p < \infty$ and in the space $C_0(\Re)$.

Proposition 3.5.12 [14]:

1- If $f \in C_0(\Re)$**,** then $\left\| p_j f \right\|_\infty \to 0$ when $j \to -\infty$.

2- If $f \in L^p(\Re)$, $1 < p < \infty$, then $\left\| p_j f \right\|_p \to 0$ when $j \to -\infty$.

Proof : we begin with $f \in C_{00}(\Re)$. If $f(x) = 0$ for $|x| > R$, we can write

$$p_{-m} f(x) = 2^{-m} C \int_{-R}^{R} f(y) K(2^{-m} x, 2^{-m} y) dy \qquad (3.5.6)$$

$$\leq 2^{-m} C \int_{-R}^{R} \left| f(y) \right| W\left(2^{-m} \frac{|x - y|}{2} \right) dy$$

hence,

$$p_{-m} f(x) \leq 2^{-m} C \left\| f \right\|_\infty 2RW(0) \to 0, \quad m \to \infty.$$

But $C_{00}(\Re)$ is dense in $C_0(\Re)$ where we have the estimate $\left\| p_j f \right\|_\infty \leq C \left\| f \right\|_\infty$.

To prove the L^p convergence, it suffices to take $f \in C_{00}(\Re)$. For $|x| \le R$ the estimate (3.5.6) shows that $\int_{-R}^{R} |p_{-m} f(x)|^p \, dx \to 0$. For $x > R$ we make the substitution $v = 2^{-m}(x - y)$ to write

$$|p_{-m} f(x)| \le \|f\|_\infty \int_{-2^{-m(x+R)}}^{2^{-m(x+R)}} |W(v)| \, dv$$

$$\le 2^{-m} \|f\|_\infty RW\left(2^{-m}(x - R)\right)$$

$$\int_{R}^{\infty} |p_{-m} f(x)|^p \, dx \le 2^{-mp} \|f\|_\infty^p \int_{R}^{\infty} \left|W\left(2^{-m}(x - R)\right)\right|^p dx$$

$$= 2^{-mp} 2^m \|f\|_\infty^p \int_0^\infty |W(y)|^p \, dy$$

$$= 2^{m(1-p)} \|f\|_\infty^p \|W\|_p^p \to 0$$

with a similar estimate for $t < -R$.

In exact parallel with the case of Haar series, the large scale projection operators to do not behave well on $L^1(\Re)$. This means that we restrict the range of *p* when formulating a general L^p convergence theorem for wavelet series. Similarly, we must restrict to $C_0(\Re)$, since the identity $p_j 1 = 1$ shows that $p_j f \to 0$ is false in general when $f \in B_{uc}(\Re)$, for $j \to -\infty$.

Combining proposition (3.5.12) and theorem (3.5.11), gives a complete picture of the convergence of one-dimensional wavelet series in the spaces $C_0(\Re)$ and $L^p(\Re)$, $1 < p < \infty$. This can be restated in a separate theorem.

Theorem 3.5.13 [6, 14]: Suppose that the scaling function φ has radial decreasing L^1-majorant *W*.

1. If $f \in C_0(\Re)$, then the sum $\sum_{j=-m}^{n} \sum_{k \in Z} \varphi_{j,k}(x) \int_{\Re} f(y) \overline{\varphi}_{j,k}(y) dy$ converges uniformly to *f* when $m, n \to \infty$.

2. If $f \in L^p(\Re), 1 < p < \infty$, then the sum $\sum_{j=-m}^{n} \sum_{k \in Z} \varphi_{j,k}(x) \int_{\Re} f(y) \overline{\varphi}_{j,k}(y) dy$ converges to *f* in $L^p(\Re)$ when $m, n \to \infty$.

Chapter four

Convergence Analysis

4.1. Introduction

4.2. Rates of decay of Fourier coefficients

4.3. Rate of convergence of Fourier series in L^2

4.4. Rates of decay of Haar coefficients

4.5. Rate of convergence of Haar series

4.6. Rate of convergence of wavelet series

4.7. Conclusion

Chapter four
Convergence Analysis

4.1 Introduction

Under certain conditions, a function can be represented with a sum of sine and cosine functions, which is called a Fourier series. This classical method is used in applications such as storage of sound waves and visual images on a computer. One problem with representing a functions with this type of series is that it takes an infinite number of terms to represent such function. In practice, only a finite number of terms can be used. Higher accuracy require the sum of more terms in this series and this will take up more computer time and storage space. A new type of sum called a wavelet series was first introduced in the 1980's and found to be more efficient, in storage and processing, than Fourier series. Efficiency of a series representation of a signal (function) depends on its convergence which in turn depends on the rate of decay in its coefficients. In this chapter, we will investigate the superiority of the wavelet series in representing signals over the Fourier series through the rate of decay of the coefficients for both Fourier and wavelet series.

4.2 Rates of decay of Fourier coefficients

The Riemann- lebesgue lemma state that the Fourier coefficients of an integrable and 2π-periodic function f vanish at infinity, but it provides no further information about the speed of convergence to zero for such function. In this section, we shall show the relationship between the smoothness of f and the magnitude of its Fourier coefficients c_n. (The smoothness of f is measured by the number of times it is differentiable).

Definition 4.2.1 [16]: Class C^K

We say that f belong to the class C^K if f is K times continuously differentiable.

Definition 4.2.2 [14]: a Holder condition

Let f be a function defined on $\Re$. We say that f satisfy a Holder condition with exponent $\alpha \in (0,1)$, if $|f(x) - f(y)| \le M|x-y|^{\alpha}$, $M > 0$.

Remark [16]:

Belonging to the class C^K or satisfying a Holder condition are two possible ways to describe the smoothness of a function.

Proposition 4.2.3 [14]: Suppose that $f \in C(\Re)$ has a modulus of continuity: $\omega(\delta) := \sup_{|x-y|\le\delta} |f(x) - f(y)|$. Then $|c_n| \le \frac{1}{2}\omega\left(\frac{\pi}{n}\right)$.

Proof: Since $e^{-inx} = -e^{-in(x-\pi/n)}$, we have

$$\int_0^{2\pi} f(x)e^{-inx}dx = -\int_0^{2\pi} f(x+\pi/n)e^{-inx}dx = \frac{1}{2}\int_0^{2\pi}[f(x) - f(x+\pi/n]e^{-inx}dx$$

and hence, $c_n = \frac{1}{4\pi}\int_0^{2\pi}[f(x) - f(x+\pi/n)]e^{-inx}dx$,

So,

$$\begin{aligned} |c_n| &\le \frac{1}{4\pi}\int_0^{2\pi}|f(x) - f(x+\pi/n)||e^{-inx}||dx| \\ &= \frac{1}{4\pi}.w(\pi/n).2\pi \\ &= \frac{1}{2}.w(\pi/n). \end{aligned}$$

Corollary 4.2.4 [14]: If f satisfies a Holder condition with exponent $\alpha \in (0,1)$, we see that $c_n = O(n^{-\alpha})$, $|n| \to \infty$.

Proof: f satisfies a Holder condition with exponent $\alpha \in (0,1)$ means: $|f(x) - f(x+h)| \le Ch^{\alpha}$, C : constant, take $h = \pi/n$, and use proposition (4.2.3) to get

$$|c_n| \le \frac{1}{4\pi}\int_0^{2\pi} |f(x) - f(x+\pi/n)||e^{-inx}||dx|$$

$$= \frac{1}{4\pi}.C(\pi/n)^{\alpha}.2\pi$$

$$= \left(\frac{\mathrm{C}\pi^{\alpha}}{2}\right)\left(\frac{1}{n}\right)^{\alpha}$$

Therefore, $c_n = O(n^{-\alpha}), |n| \to \infty$.

If we want to obtain a more precise estimation, we can assume that f is absolutely continuous as follows:

Proposition 4.2.5 [14]: If $f \in C^K(\Re), K \ge 1$ are absolutely continuous. Then $c_n = o\left(1/|n|^K\right), \quad |n| \to \infty$

Proof: Assume $f \in C^K(\Re)$**,** $K \ge 1$. Then

$$c_n' = \frac{1}{2\pi}\int_0^{2\pi} f'(x)e^{-inx}dx$$

$$= \frac{1}{2\pi} f(x)e^{-inx}\Big|_0^{2\pi} + in\frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-inx}dx$$

$$= inc_n,$$

which by induction yields

$$c_n = \frac{c_n^K}{(in)^K}. \tag{4.2.1}$$

Since $f^K \in L^1(\Re)$, we have $\lim\limits_{|n|\to\infty} c_n^K = 0$, which implies $\lim\limits_{|n|\to\infty}|n^K c_n| = \lim\limits_{|n|\to\infty} c_n^K = 0$. So $c_n = o\left(1/|n|^K\right), |n| \to \infty$.

Corollary 4.2.6 [14]: If $f^{(K)}$ satisfies a Holder condition with exponent $\alpha \in (0,1)$, then $c_n = O(n^{-\alpha-K}), |n| \to \infty$.

Proof: By (4.2.1)

$$(in)^K c_n = c_n{}^K = \frac{1}{4\pi}\int_0^{2\pi}\left[f^K(x) - f^K(x+\pi/n)\right].e^{-inx}dx$$

$$\Rightarrow \qquad |c_n| \le \frac{1}{4\pi|n|^K}\int_0^{2\pi}\left|f^K(x) - f^K(x+\pi/n)\right|.\left|e^{-inx}\right||dx|$$

$$= \frac{1}{4\pi|n|^K}.C\left(\frac{\pi}{n}\right)^{\alpha}.2\pi$$

$$= \left(\frac{C.\pi^{\alpha}}{2}\right).\frac{1}{n^K.n^{\alpha}} = M.(n)^{-\alpha-K} \quad .$$

Note that the smoothness of f is directly related to the decay of the Fourier coefficients, and in general, the smoother of the function, the faster decay. As a result, we can expect that relatively smooth functions equal their Fourier series.

4.3 Rate of convergence of Fourier series in L^2

Definition 4.3.1: Mean square error

The mean square error $\|S_N f - f\|_2^2$ of the Fourier series of f is defined by:

$$\|S_N f - f\|_2^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi}|f - S_N f|^2\,dx$$

Where $S_N f$ is the N^{th} partial sum of f.

Parseval's theorem allows us to reduce the study of rate of convergence to the estimation of series. The N^{th} Fourier coefficient of $S_N f - f$ is zero for $|n| \le N$, therefore $S_N f - f = \sum_{|n|>N} c_n e^{inx}$, hence by Parseval's theorem we have

$$\|S_N f - f\|_2^2 = \sum_{|n|>N}|c_n|^2 \qquad (4.3.1)$$

This can be used to estimate the mean square error in terms of the smoothness of f.

Proposition 4.3.2 [14]: Suppose that $f \in C^K(\Re)$, then $c_n = O\left(1/|n|^K\right), |n| \to \infty$ and $\|S_N f - f\|_2^2 = C\sum_{n>N} n^{-2K} = O\left(N^{1-2K}\right), \quad N \to \infty$.

Which gives an upper bound for the mean square error when $N \to \infty$.

Proof: By (4.2.1), we have

$$c_n = \frac{c_n^K}{(in)^K} = \frac{1}{(in)^K 2\pi}\int_0^{2\pi} f^K(x)e^{-inx}dx$$

$$|c_n| \le \frac{1}{2\pi|n^K|}\int_0^{2\pi}|f^K||dx| \le \frac{M}{|n^K|}$$, and hence $c_n = O\left(1/|n|^K\right), |n| \to \infty$.

Now by (4.3.1),

$$\|S_N f - f\|_2^2 = \sum_{|n|>N}|c_n|^2 = \sum_{|n|>N}\frac{M^2}{n^{2K}} = C\sum_{|n|>N} n^{-2K}$$

Example 4.3.1: Consider the function $f(x) = x^3 - \pi^2 x$, over $[-\pi, \pi]$.The Fourier series of f is defined by: $f(x) = 12\sum_{n=1}^{\infty}\frac{(-1)^n}{n^3}\sin nx$.

So by (4.3.1), we find that $\|S_N f - f\|_2^2 = \sum_{|n|>N}\frac{144}{n^6}$

Example 4.3.2: Let $f(x) = x^2$, over$[-\pi, \pi]$, then the Fourier series of f is defined by: $f(x) = \frac{\pi^2}{3} + \sum_{n=1}^{\infty}\frac{4(-1)^n}{n^2}\cos nx$. So

$$\|S_N f - f\|_2^2 = \sum_{|n|>N}\frac{16}{n^4}$$

4.4 Rates of decay of Haar coefficients

We have seen that the smoothness of the function is reflected in the decay of its Fourier coefficients. Specifically, if f is periodic and C^K on $\Re$, then $c_n = O\left(1/|n|^K\right), |n| \to \infty$. Where c_n is the Fourier coefficient of f.

This can be regarded as a statement about the frequency content of smooth functions, namely that smoother functions tend to have smaller high frequency components than do functions that are not smooth.

However, no such estimate holds for the Haar series. To see this, simply note that the function $f(x) = e^{ix}$ is periodic and is C^∞ on $\Re$ with all of its derivatives bounded by 1. But have

$$\left| < f, \psi_{j,k} > \right| = 2^{-j/2} \frac{\sin^2\left(\left(1/4\right)2^{-j}\right)}{\left(\left(1/4\right)2^{-j}\right)}$$

and since $\sin\left(\left(1/4\right)2^{-j}\right) \approx \left(\left(1/4\right)2^{-j}\right)$ for large *j*. this means that $\left| < f, \psi_{j,k} > \right| \approx 2^{-3j/2}.\frac{1}{4}$ for large *j*. But this is the same rate of decay as we will see later for functions continuous but with a discontinuous first derivative. Hence, the smoothness of a function does not affect the rate of decay of its Haar coefficients.

Proposition 4.4.1: If f satisfies a Holder condition with exponent $\alpha \in (0,1]$, then
$< f, \psi_{j,k} > = O\left(2^{-j(\alpha+1/2)}\right)$

Haar Coefficients near Jump Discontinuity [1]

Suppose that f is a function defined on $[0,1]$, with a jump discontinuity at $x_0 \in (0,1)$ and continuous at all other points in $[0,1]$. Here we analyze the behavior of Haar coefficients when x_0 is inside or outside the dyadic interval $I_{j,k}$. In particular, we can find the location of a jump discontinuity just by examining the absolute value of the Haar coefficients.

For simplicity, let us assume that f is C^2 on $[0, x_0]$ and $[x_0,1]$. This means that both f' and f'' exist, and continuous functions, and hence bounded on each of these intervals. For fixed $j \geq 0$ and $0 \leq k \leq 2^j - 1$, and let $x_{j,k}$ be the mid point of the interval $I_{j,k}$; that is, $x_{j,k} = \left(k + \frac{1}{2}\right)2^{-j}$. There are now two possibilities, either $x_0 \in I_{j,k}$ or $x_0 \notin I_{j,k}$.

Case 1: If $x_0 \notin I_{j,k}$, then for large *j*,

$$\left| < f, \psi_{j,k} > \right| \approx \left(\frac{1}{4}\right) 2^{-3j/2} \left| f'(x_{j,k}) \right| = O\left(2^{-3j/2}\right)$$

Proof: If $x_0 \notin I_{j,k}$, then expanding $f(x)$ about $x_{j,k}$ by Taylor's formula

$$f(x) = f(x_{j,k}) + f'(x_{j,k})(x - x_{j,k}) + \frac{1}{2} f''(\xi_{j,k})(x - x_{j,k})^2$$

where $\xi_{j,k} \in I_{j,k}$. Now using the fact that $\int \psi_{j,k}(x)\,dx = 0$,

$$< f, \psi_{j,k} > = \int_{I_{j,k}} f(x)\psi_{j,k}(x)\,dx$$

$$= f(x_{j,k}) \int_{I_{j,k}} \psi_{j,k}(x)\,dx + f'(x_{j,k}) \int_{I_{j,k}} \psi_{j,k}(x)\left(x - x_{j,k}\right)dx$$

$$+ \frac{1}{2} \int f''(\xi_{j,k})\left(x - x_{j,k}\right)^2 \psi_{j,k}(x)\,dx$$

$$= f'(x_{j,k}) \int_{I_{j,k}} x\psi_{j,k}(x)\,dx + r_{j,k}(x) \tag{4.4.1}$$

where $\left|r_{j,k}(x)\right| = \frac{1}{2}\left|\int_{I_{j,k}} f''(\xi_{j,k})\left(x - x_{j,k}\right)^2 \psi_{j,k}(x)\,dx\right|$.

Now

$$\int_{I_{j,k}} x\psi_{j,k}(x)\,dx = \int_{k2^{-j}}^{(k+1/2)2^{-j}} 2^{j/2}\,x\,dx - \int_{(k+1/2)2^{-j}}^{(k+1)2^{-j}} 2^{j/2}\,x\,dx$$

$$= 2^{j/2}\left\{\left[\frac{x^2}{2}\right]_{k2^{-j}}^{(k+1/2)2^{-j}} - \left[\frac{x^2}{2}\right]_{(k+1/2)2^{-j}}^{(k+1)2^{-j}}\right\}$$

$$= 2^{j/2}.2^{-2j}.\frac{1}{2}\left\{\left(k + \frac{1}{2}\right)^2 - k^2 - (k+1)^2 + \left(k + \frac{1}{2}\right)^2\right\}$$

$$= -\frac{1}{4}2^{-3j/2}. \tag{4.4.2}$$

From (4.4.1) and (4.4.2)

$$< f, \psi_{j,k} > = -\frac{1}{4}2^{-3j/2} f'(x_{j,k}) + r_{j,k}(x).$$

Now

$$\left|r_{j,k}(x)\right| \le \frac{1}{2}\max_{x \in I_{j,k}}\left|f''(x)\right| \int_{I_{j,k}} \left(x - x_{j,k}\right)^2 \left|\psi_{j,k}(x)\right| dx$$

$$\le \frac{2^{j/2}}{2}\max_{x \in I_{j,k}}\left|f''(x)\right| \int_{k2^{-j}}^{(k+1/2)2^{-j}} \left(x - x_{j,k}\right)^2 dx$$

$$= \frac{2^{j/2}}{2}.\frac{2^{-3j}}{3.4}.\max_{x \in I_{j,k}}\left|f''(x)\right|$$

$$= \frac{1}{24}.2^{-5j/2}.\max_{x \in I_{j,k}}\left|f''(x)\right|.$$

For large *j*, $2^{-5j/2}$ is very small compared with $2^{-3j/2}$. So

$$\left| < f, \psi_{j,k} > \right| \approx \left(\frac{1}{4}\right) 2^{-3j/2} \left| f'(x_{j,k}) \right| = O\left(2^{-3j/2}\right) \tag{4.4.3}$$

Case 2: If $x_0 \in I_{j,k}$, then for large *j*,

$$\left| < f, \psi_{j,k} > \right| \approx \left(\frac{1}{4}\right) 2^{-j/2} \left| f(x_0^-) - f(x_0^+) \right| = O\left(2^{-j/2}\right)$$

Proof: If $x_0 \in I_{j,k}$, then either it is in $I_{j,k}^l$ or in $I_{j,k}^r$. We assume that $x_0 \in I_{j,k}^l$, and the other case is similar. Now expanding $f(x)$ about x_0 by Taylor's formula, we have

$$f(x) = f(x_0^-) + f'(\xi^-)(x - x_0), \quad x \in [0, x_0), \ \ \xi^- \in [x, x_0]$$

$$f(x) = f(x_0^+) + f'(\xi^+)(x - x_0), \quad x \in [x_0, 1), \ \ \xi^+ \in [x_0, x].$$

Therefore

$$\begin{aligned} < f, \psi_{j,k} > &= \int_{I_{j,k}} f(x) \psi_{j,k}(x)\, dx \\ &= \int_{k2^{-j}}^{x_0} 2^{j/2} f\left(x_0^-\right) dx + \int_{x_0}^{(k+1/2)2^{-j}} 2^{j/2} f\left(x_0^+\right) dx - \int_{(k+1/2)2^{-j}}^{(k+1)2^{-j}} 2^{j/2} f\left(x_0^+\right) dx + \varepsilon_{j,k} \\ &= 2^{j/2}\left(x_0 - k2^{-j}\right)\left(f(x_0^-) - f(x_0^+)\right) + \varepsilon_{j,k}, \end{aligned} \tag{4.4.4}$$

where

$$\varepsilon_{j,k} = \int_{k2^{-j}}^{x_0} f'\left(\xi^-\right)(x - x_0)\psi_{j,k}\, dx + \int_{x_0}^{(k+1)2^{-j}} f'\left(\xi^+\right)(x - x_0)\psi_{j,k}\, dx.$$

Thus

$$\begin{aligned} \left|\varepsilon_{j,k}\right| &\le \max_{x \in I_{j,k} \setminus \{x_0\}} \left| f'(x) \right| \int_{I_{j,k}} \left| x - x_0 \right| \left| \psi_{j,k}(x) \right| dx \\ &\le 2^{j/2} \max_{x \in I_{j,k} \setminus \{x_0\}} \left| f'(x) \right| \int_{I_{j,k}} \left| x - x_0 \right| dx \\ &\le 2^{j/2} \max_{x \in I_{j,k} \setminus \{x_0\}} \left| f'(x) \right| . \frac{1}{4} 2^{-2j} \\ &= \frac{1}{4} \max_{x \in I_{j,k} \setminus \{x_0\}} \left| f'(x) \right| . \ 2^{-3j/2}. \end{aligned}$$

For large *j*, $2^{-3j/2}$ is very small compared with $2^{-j/2}$. So

$$\left|< f,\psi_{j,k} >\right| \approx 2^{j/2}\left|x_0 - k2^{-j}\right|\left|f(x_0^-) - f(x_0^+)\right| .$$

The quantity $\left|x_0 - k2^{-j}\right|$ is very small if x_0 is close to the left end point of $I_{j,k}^l$ and can even be zero. However, we can expect that in most cases, x_0 will be in the middle of $I_{j,k}^l$ so that $\left|x_0 - k2^{-j}\right| \approx \frac{1}{4}.2^{-j}$. Thus for large *j*,

$$\left|< f,\psi_{j,k} >\right| \approx \left(\frac{1}{4}\right)2^{-j/2}\left|f(x_0^-) - f(x_0^+)\right| = O\left(2^{-j/2}\right) \qquad (4.4.5)$$

Comparing (4.4.3) and (4.4.5), we see that the decay of $\left|< f,\psi_{j,k} >\right|$ for large *j* is considerably slower if $x_0 \in I_{j,k}$ than if $x_0 \notin I_{j,k}$.

4.5 Rate of convergence of Haar series

Proposition 4.5.1 [25]: Let *f* be continuous in $L^p(\Re)$, $1 < p < \infty$ and the partial sum of the Haar wavelet series is

$$f_N = \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1} < f,\psi_{j,k} > \psi_{j,k}$$

where $N = 2^J$ for some $J \in \mathrm{N}$. Then the error of approximation in $L^p(\Re)$ is defined by: $\left\|f - f_N\right\|_p = O\left(2^{-J/2}\right)$.

As special case for *p* = 2, the mean square error is $\left\|f - f_N\right\|_2^2 = O\left(2^{-J}\right)$.

Proof: The error of approximation in $L^p(\Re)$ is

$$\begin{aligned}\left\|f - f_N\right\|_p &= \left\|f - \sum_{j=0}^{J-1}\sum_{k=0}^{2^j-1} < f,\psi_{j,k} > \psi_{j,k}\right\|_p \\ &= \left\|\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1} < f,\psi_{j,k} > \psi_{j,k}\right\|_p \\ &= \left(\sum_{j=J}^{\infty}\sum_{k=0}^{2^j-1}\left|< f,\psi_{j,k} >\right|^p\right)^{1/p} \\ &\sim \left(\sum_{j=J}^{\infty} 2^{-jp/2}\right)^{1/p} \sim \left(2^{-J/2}\right) = O\left(2^{-J/2}\right).\end{aligned}$$

4.6 Rate of convergence of wavelet series

In this section we will examine how well a function f can be approximated pointwise by wavelets in V_j, as well as approximated in the L^2 sense. We will also look at the rate of decay of the wavelet coefficients $< f, \psi_{j,k} >$ as $j \to \infty$.

Let's start with poitwise convergence. Fix $j = J$ and suppose that $f \in C^K(T)$, where T: is the neighborhood $|x - x_0| \le \frac{1}{2^J}$ of x_0. We want to estimate the pointwise error $|f - p_J f|$ in T.

Proposition 4.6.1 [17]: Suppose that $f \in C^K(T)$**,** and $|f^K|$ has upper bound M_K in T, then

$$|f - p_J f| \le \frac{CM_K}{2^{J(K+1)}} = O\left(2^{-J(K+1)}\right)$$

where C is a constant, independent of f and J.

Proof: The proof can be found in [17]

Note that this is a local estimate; it depends on the smoothness of f in T. Thus once the wavelets is fixed, the local rate of convergence depending only on the local behavior of f. This is different from the Fourier series or Fourier integrals where a discontinuity of a function at one point can slow the rate of convergence at all points.

Now we turn to the estimation of the wavelet expansion coefficients $< f, \psi_{j,k} >$. Recall that any wavelet $\psi(x)$ that comes from an MRA must satisfy $\int_{\Re} \psi(x)dx = 0$, and we say that the *zeroth* moment of $\psi(x)$ is vanishing, so if the integral $\int_{\Re} x^K \psi(x)dx = 0$, we say that the *Kth* moment of $\psi(x)$ is vanishing. We will see that vanishing moment have results for the efficient representation of functions. Specifically we will see that the wavelet series of a smooth function will converge very rapidly to the function as long as the wavelet has a lot of vanishing moments. This means that in this case, relatively few wavelet coefficients will be required in order to get a good approximation. Now we will show that the wavelet coefficients of such functions will have rapid decay as $j \to \infty$. To make the proof easier, we will assume that $\psi(x)$ has compact support.

Proposition 4.6.2 [1]: Suppose that $f \in C^K(\Re), K \in \mathrm{N}$**,** and $\left|f^{(K)}(x)\right|$ has a uniform upper bound M_K on $\Re$. Assume that the function $\psi(x)$ has K vanishing moment with compact support, and $\int_{\Re} \left|\psi_{j,k}(x)\right|^2 dx = 1, \forall j,k \in \mathrm{Z}$,then we have the estimate

$$\left|< f, \psi_{j,k} >\right| \le \frac{CM_K}{2^{j(K+1/2)}} = O\left(2^{-j(K+1/2)}\right)$$

where C is a constant, independent of f, j, k.

Proof: Suppose that $\psi(x)$ is supported in the interval I, which has the form $I_{0,0} = [0, a]$ for some $a > 0$. It follows that the function $\psi_{j,k}(x)$ is supported in the interval $I_{j,k} = \left[2^{-j}k, 2^{-j}(k+a)\right]$, and $\left|I_{j,k}\right| = 2^{-j}a$. Now let $x_{j,k} = 2^{-(j+1)}a + 2^{-j}k$ be the center of the interval $I_{j,k}$.

Since $f \in C^K(\Re)$, for each $j,k \in \mathrm{Z}$, $f(x)$ can be expanded in a Taylor expansion about the point $x_{j,k}$. That is,

$$f(x) = f(x_{j,k}) + (x - x_{j,k}) f'(x_{j,k}) + \ldots + \frac{1}{(K-1)!}(x - x_{j,k})^{K-1} f^{(K-1)}(x_{j,k}) + R_K(x),$$

where

$$R_K(x) = \frac{1}{K!}(x - x_{j,k})^K f^{(K)}(\xi)$$

for some ξ between $x_{j,k}$ and x. If $x \in I_{j,k}$, then we have the estimate

$$\left|R_K(x)\right| \le \frac{1}{K!} 2^{-K(j+1)} a \max_{x \in I_{j,k}} \left| f^{(K)}(x) \right|. \tag{4.6.1}$$

Now we can compute the wavelet coefficients as follows:

$$\begin{aligned}
< f, \psi_{j,k} > &= \int_{\Re} f(x) \overline{\psi_{j,k}(x)}\, dx \\
&= \int_{\Re} \left(\sum_{l=0}^{K-1} \frac{1}{l!} (x - x_{j,k})^l f^{(l)}(x_{j,k}) + R_K(x) \right) \overline{\psi_{j,k}(x)}\, dx \\
&= \left(\sum_{l=0}^{K-1} \frac{1}{l!} \int_{\Re} (x - x_{j,k})^l \overline{\psi_{j,k}(x)}\, dx \right) + \int_{\Re} R_K(x) \overline{\psi_{j,k}(x)}\, dx \\
&= \int_{I_{j,k}} R_K(x) \overline{\psi_{j,k}(x)}\, dx.
\end{aligned}$$

Now applying the estimate (4.6.1) and the Cauchy-Schwarz inequality,

$$\begin{aligned}
\left| < f, \psi_{j,k} > \right| &= \left| \int_{I_{j,k}} R_K(x) \overline{\psi_{j,k}(x)}\, dx \right| \\
&\le \frac{1}{K!} 2^{-K(j+1)} a \max_{x \in I_{j,k}} \left| f^{(K)}(x) \right| \int_{I_{j,k}} \left| \psi_{j,k}(x) \right| dx \\
&\le \frac{1}{K!} 2^{-K(j+1)} a \max_{x \in I_{j,k}} \left| f^{(K)}(x) \right| \left| I_{j,k} \right|^{1/2} \left(\int_{I_{j,k}} \left| \psi_{j,k}(x) \right|^2 dx \right)^{1/2} \\
&= \frac{1}{K!} 2^{-K(j+1)} a \max_{x \in I_{j,k}} \left| f^{(K)}(x) \right| . 2^{-j/2} a^{1/2} \\
&= 2^{-j(K+1/2)} \left(\frac{1}{K!} \max_{x \in I_{j,k}} \left| f^{(K)}(x) \right| . a^{3/2} 2^{-K} \right) \\
&\le 2^{-j(K+1/2)} \left(\frac{1}{K!} a^{3/2} 2^{-K} M_K \right)
\end{aligned}$$

Note that with $C = \frac{1}{K!} a^{3/2} 2^{-K}$, the proof is complete.

We already know that the wavelet basis is complete in $L^2[-\infty, \infty]$. Let consider the decomposition: $L^2[-\infty, \infty] = V_J \oplus \sum_{j=J}^{\infty} W_j$.

We want to estimate the L^2 error $\| f - p_J f \|_2^2$ as follows:

Proposition 4.6.3: Suppose that $f \in C^K(\Re)$ and has bounded support say the interval $(0, a)$, if $\left| f^{(K)}(x) \right|$ has a uniform upper bound M_K then

$$\|f - p_J f\|_2^2 = \sum_{j=J}^{\infty}\sum_{k=0}^{2^j a}\left|< f,\psi_{j,k} >\right|^2 < \frac{2C^2 M_K^2 a}{2^{2KJ}} = O\left(2^{1-2KJ}\right).$$

Proof: The proof is easy by using Proposition 4.6.2.

Jackson's approximation theorem

Now we formulate results that relate the speed of convergence of wavelet series to the smoothness of f. We focus attention on the rate of decay of $\|f - p_j f\|_p$.

In order to measure the smoothness of a function, we introduce the L^p modulus of continuity:

$$\omega_p(f;\delta) = \sup_{0<h<\delta}\|f(x) - f(x-h)\|_p$$

This is defined if $f \in L^p(\Re)$ or not.

Proposition 4.6.4 [14, 18]: The L^p modulus of continuity satisfies the following conditions:

a. $\delta \to \omega_p(f;\delta)$ is monotone increasing.

b. If $f \in L^p(\Re)$, then $\omega_p(f;\delta) \to 0$ as $\delta \to 0$.

c. $\omega_p(f;\delta_1 + \delta_2) \le \omega_p(f;\delta_1) + \omega_p(f;\delta_2)$.

d. $\omega_p(f_1 + f_2;\delta) \le \omega_p(f_1;\delta) + \omega_p(f_2;\delta)$.

e. If $\omega_p(f;\delta) < \infty,\ \forall\ \delta > 0$, then $|f|^p \in L^1_{\text{loc}}(\Re)$.

f. $\omega_p(f;m\delta) \le m\omega_p(f;\delta)$.

Proof: The proof can be found in [14].

In order to prove suitable approximation theorems, we need to consider a small class of scaling functions, defined by an estimate of the form

$$|\varphi(x)| \le \frac{A}{\left(1+|x|\right)^B},\ B > 2. \tag{4.6.1}$$

Lemma 4.6.5 [14]: If φ satisfies (4.6.1), then the wavelet kernel $K(x,y)$ satisfies the estimate

$$|K(x,y)| \le \frac{A}{(1+|x-y|)^B} \tag{4.6.2}$$

The direct approximation (Jackson's estimate) is the following statement.

Theorem 4.6.6 [14, 18]: Jackson's inequality

Suppose that the scaling function satisfies (4.6.1). Then there exist a constant *C* such that for all $f \in MC_p(\Re)$

$$\|f - p_j f\|_p \le C\omega_p\left(f;2^{-j}\right) \tag{4.6.3}$$

where the space $MC_p(\Re)$ is defined by:=$\{ f : \omega_p(f;\delta) < \infty \quad \text{for all } \delta > 0 \}$.

Note that we do not assume that $f \in L^p(\Re)$.

We can reduce this to study of p_0 by introducing the dilation operator.

Definition 4.6.7 [14]: Dilation operator

The dilation operator defined by: $J_r f(x) = f(2^r x),\ r \in Z$.And satisfies the following properties:

a. Commutation relation: $p_j J_r = J_r p_{j-r}$.
b. Norm relation: $\|J_j f\|_p = 2^{-j/p}\|f\|_p$.
c. $\omega_p(J_a f;\delta) = 2^{-a/p}\omega_p(f;2^a\delta)$.

Now we return to proof theorem (4.6.6).

Proof: Suppose we have (4.6.3) for $j=0$ with some constant *C*. Then by using the properties a, b and c in definition (4.6.7) we get

$$\begin{aligned}\|p_j f - f\|_p &= \|J_j p_0 J_{-j} f - J_j J_{-j} f\|_p \\ &= 2^{-j/p}\|p_0 J_{-j} f - J_{-j} f\|_p \\ &\le \mathrm{C}\, 2^{-j/p}\omega_p(J_{-j} f;1) = C\omega_p(f;2^{-j}).\end{aligned}$$

So it suffices to consider $j=0$. From (3.5.4) using property (3) for the wavelet kernel we can write

$$f(x)-p_0 f(x)=\int_{-\infty}^{\infty}[f(x)-f(y)]K(x,y)\,dy$$

From (4.6.2) we get

$$\|f-p_0 f\|_p^p=\int_{-\infty}^{\infty}\left|\int_{-\infty}^{\infty}[f(x)-f(y)]K(x,y)\,dy\right|^p dx$$

$$\le A\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\frac{|f(x)-f(y)|\,dy}{(1+|y-x|)^B}\right)^p dx$$

$$=A\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}\frac{|f(x)-f(x+u)|\,du}{(1+|u|)^B}\right)^p dx$$

We pick $a>0, b>0$ so that $B=a+b$ and $ap>p+1, bp'>1$ (where as $p^{-1}+p'^{-1}=1$)and applying Holder's inequality to the inside integral we get

$$\|f-p_0 f\|_p^p\le A\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|f(x)-f(x+u)|^p\,du}{(1+|u|)^{ap}(1+|u|)^{bp}}\,dx$$

$$\le A\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{|f(x)-f(x+u)|^p\,du}{(1+|u|)^{ap}}\left(\int_{-\infty}^{\infty}\frac{du}{(1+|u|)^{bp'}}\right)^{p/p'}dx$$

$$\le A\int_{-\infty}^{\infty}\frac{\omega_p(f;|u|)^p\,du}{(1+|u|)^{ap}}$$

We divide the last integral in to two parts and estimate each part separately as follows:

$$\int_{-1}^{1}\frac{\omega_p(f;|u|)^p\,du}{(1+|u|)^{ap}}\le C\omega_p(f;1)^p\ ,$$

and

$$\int_{-\infty}^{1}+\int_{1}^{\infty}\frac{\omega_p(f;|u|)^p\,du}{(1+|u|)^{ap}}\le 2\int_{1}^{\infty}\frac{\omega_p(f;|u|)^p\,du}{(1+|u|)^{ap}}$$

using proposition (4.6.4,f) to get

$$\int_{-\infty}^{1} + \int_{1}^{\infty} \frac{\omega_p(f;|u|)^p\,du}{(1+|u|)^{ap}} \leq C \int_{1}^{\infty} \frac{u^p \omega_p(f;1)^p\,du}{(1+u)^{ap}}$$

$$\leq C\,\omega_p(f;1)^p \int_{1}^{\infty} \frac{u^p\,du}{(1+u)^{ap}}$$

$$\leq C\omega_p(f;1)^p$$

since $ap > p+1$.

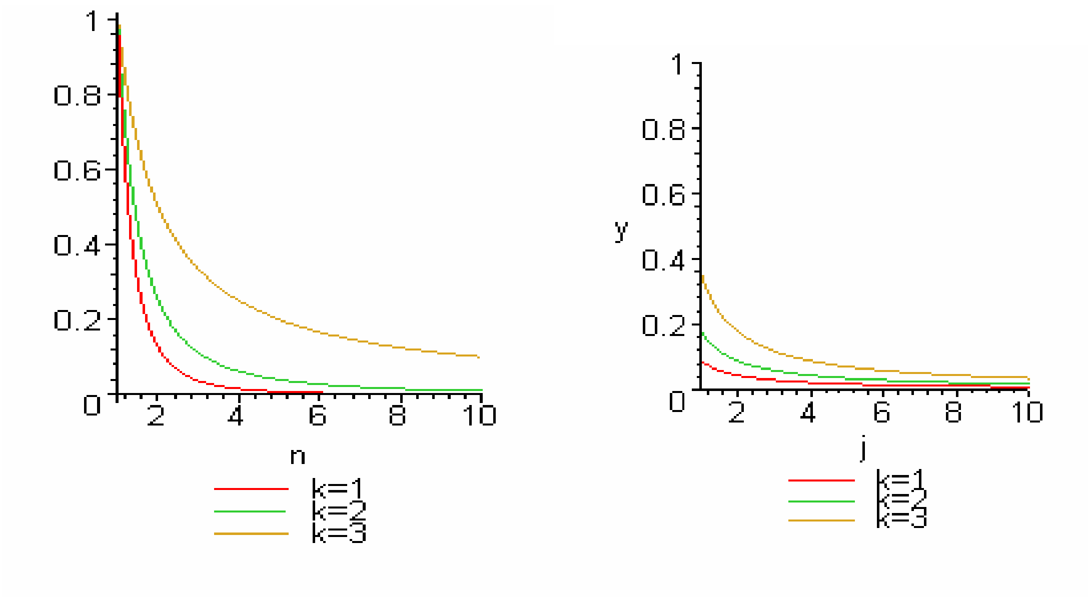
Corollary 4.6.8 [14]: If f satisfies a Holder condition with exponent $\alpha \in (0,1]$, then $\|f - p_j f\|_p \leq C 2^{-j\alpha} = O(2^{-j\alpha})$.

4.7 Conclusion

We can summarize the results we obtained in this chapter as follows:

1. If a function f is sufficiently smooth; i.e. $f \in C^K(\Re)$, then the rate of decay of the Fourier coefficients of f is of order $O(n^{-K})$ with mean square error of order $O(N^{1-2K})$, whereas the rate of decay of its wavelet coefficients is of order $O(2^{-j(K+1/2)})$ with mean square error of order $O(2^{1-2KJ})$.
2. If a function f is satisfies a Holder condition with exponent $\alpha \in (0,1]$, then the rate of decay of the Fourier coefficients of f is of order $O(n^{-\alpha})$, whereas the rate of decay of its wavelet coefficients is of order $O(2^{-j(\alpha+1/2)})$.

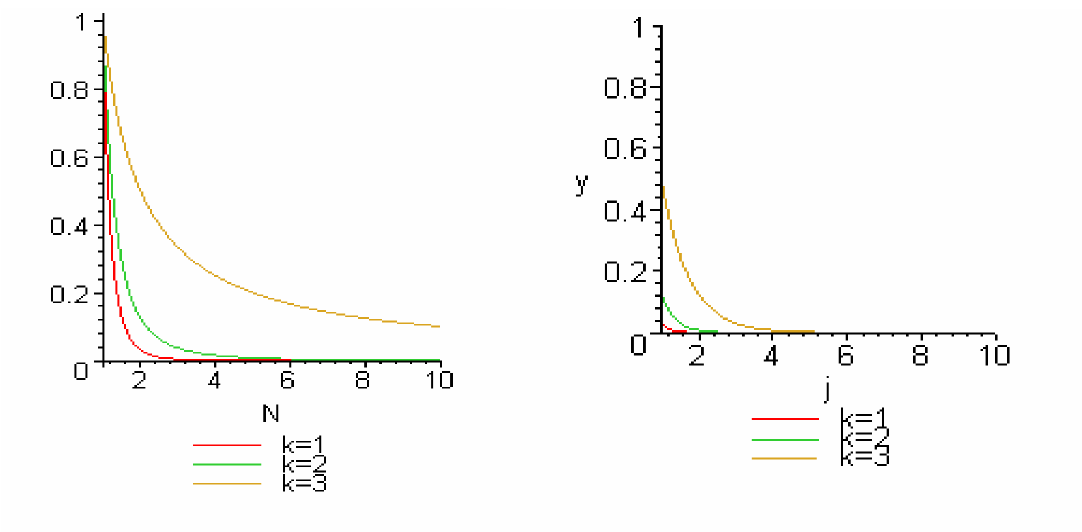
Note that the smoothness of f is directly related to the rate of decay for both coefficients; Fourier and wavelet, but does not affect the rate of decay of the Haar coefficients. See figure (1).



(a) **(b)**

Figure 1

From the above results we expect that under the same condition of f the speed of convergence of wavelet series is faster than the speed of convergence of its Fourier series, and this is one advantage for wavelet. See figure (2).



(a) **(b)**

Figure 2

Finally, I will end this thesis by setting some differences between both Fourier and wavelets transform.

1. As we show in chapter 2, the Fourier series of a function with a jump discontinuity exhibits Gibb's phenomenon. That is, the partial sums overshoot the function near the discontinuity and this overshoot continues no matter how many terms are taken in the partial sum. Gibb's phenomenon does not occur if the partial sum replaced by the arithmetic mean σ_N. Since the wavelet expansions have convergence properties similar to σ_N, we might expect them not to exhibit Gibb's phenomenon.
2. We can see that unlike the trigonometric system the Haar system provide the uniform convergence on the partial sums for continuous function on $[0,1]$. This property is shared by most wavelets in contrast with the Fourier basis for which the best we can expect for continuous functions is pointwise convergence a.e. Also, the partial sums of the Fourier series of continuous functions do not necessarily converge. To expect the uniform convergence we assume that f is a piecewise smooth function.
3. The wavelet coefficients in the wavelet series expansion of a function are the integral wavelet transform of the function evaluated at certain dyadic points $\left(\frac{k}{2^j},\frac{1}{2^j}\right)$. No such relationship exists between Fourier series and Fourier transform, which are applicable to different classes of functions; Fourier series applies to functions that are square integrable in $[0,2\pi]$, whereas Fourier transform is for functions that are in $L^2(\Re)$. Both wavelet series and wavelet transform are applicable to functions in $L^2(\Re)$

References

1. Walnut D. **An introduction to Wavelets Analyses**. Boston: Birkhauser Basel; 2002. 449pp.
2. Daubechies I. **Ten Lecture on Wavelet**. Society for Industrial and Applied Mathematics; 1992. 343PP.
3. Chui C. **An introduction to Wavelets**. New York: Academic Press; 1992. 366PP.
4. Boggess A.L, Narcowich F. J. **A First Course in Wavelets with Fourier Analysis**. 292pp.
5. Zygmund A. **Trigonometric Series**. Poland: Warsaw; 1935. 331PP.
6. Hernandez Eu, Weiss Gu. **A First Course on Wavelets**. New York: CRC Press; 1996. 493pp.
7. Mallat S. **A wavelet Tour of Signal Processing**. 2^{nd} edition. London: academic Press; 1998. 668pp.
8. Allen R.L, Mills D.W. **Signal Analysis: Time, Frequency, Scale, and Structure**. Canada: IEEE Press; 2004. 962pp.
9. Burrus S.C, Gopinath R.A, Guo H. **Introduction to Wavelets and Wavelet Transforms**. New Jersey: Prentice- Hall, Inc; 1998. 281pp.
10. Hanna J. R, Rowland J. H. **Fourier series, Transforms, and Boundary Value Problems**. 2^{nd} edition. New York: John Wiley & Sons. Inc; 1990. 365pp.
11. Kahane J-P, Lemmarie P-G. **Fourier series and Wavelets**. Amsterdam: OPA (Overseas Publishers Association); 1995. 394PP.

12. Debnath L. **Wavelet transform and their applications**. Boston: Birkhauser; 2002. 575pp.

13. Hong Do, Wang Ji. **Real Analysis with an Introduction to Wavelets**. London: academic Press .342pp.

14. Pinsky M**. Introduction to Fourier Analysis and Wavelets**. United State of America: Wadsworth Group; 2002. 387pp.

15. Goswami J.C, Chan A.K. **Fundamentals of Wavelets: Theory, Algorithm, and Applications**. New York: John Wiley & Sons. Inc. 319pp.

16. Stein E. M, Shakarchi Ra. **Fourier Analysis: An Introduction**. New Jersey: Princeton University Press; 2002. 329pp.

17. Miller W. **Introduction to the mathematics of wavelets**. 2006. 294pp.

18. Wojtaszczyk P. **A Mathematical introduction to wavelets**. Cambridge: Cambridge University press; 1997. 273pp.

19. Bruckner A. M, Bruckner J. B., Thomson B. S**. Real Analysis**. New Jersey: Prentice- Hall, Inc; 1997. 682pp.

20. Walker J. **Fourier Analysis**. Oxford University Press; 1988.

21. Boyce W. E, Diprima R. C. **Elementary Differential Equations & Boundary Value Problem**. 7^{th} edition. New York: John Wiley & sons, Inc; 2001. 1310pp.

22. Folland G. **From calculus to wavelets: A New Mathematical Technique, wavelet analysis**. University of Washington; 1997.

23. Roland N. **Fourier & Wavelet Representations of Functions**. Furman University, Electronic Journal of under graduate mathematics; 2000.

24. Walker J. **Fourier series**. Academic Press, Encyclopedia of physical science & Technology.

25. Islam M. **Approximation by wavelets**. Saudi Arabia: King FAHD University of PETROLEUM & Minerals; 2005. 85pp.

Appendix

Basic Theorems

Theorem 1: Cauchy-Schawrz Inequality

Let $f(x)$ and $g(x)$ be L^2 on the interval *I*, then

$$\left|\int_I f(x)\,g(x)\,dx\right| \le \left(\int_I |f(x)|^2\,dx\right)^{1/2} \left(\int_I |g(x)|^2\,dx\right)^{1/2}$$

Theorem 2: Holder Inequality

If p and q are non negative real numbers such that $\frac{1}{p}+\frac{1}{q}=1$, and if $f \in L^p$ and $g \in L^q$, then $f\,g \in L^1$ and $\|f\,g\|_1 \le \|f\|_p\,\|g\|_q$.

Theorem 3: Dominated convergence theorem

Suppose $f_n(x) \to f(x)$ almost everywhere. If $|f_n(x)| \le g(x)$ for all *n*, and $\int g(x)dx < \infty$, then *f* is integrable, and $\int f(x)dx = \lim_{n\to\infty} \int f_n(x)dx$.

Theorem 4: Taylor's Theorem

Suppose that $f(x)$ is *n*-times continuously differentiable on some interval *I* containing the point x_0. Then for $x \in I$, $f(x)$ can be written

$$f(x) = f(x_0) + (x-x_0)f'(x_0) + \frac{(x-x_0)^2}{2} f''(x_0) + \ldots + \frac{(x-x_0)^{n-1}}{(n-1)!} f^{(n-1)}(x_0) + \frac{(x-x_0)^n}{n!} f^{(n)}(\xi)$$

where ξ is some point between x_0 and *x*.

Theorem 5: Minkowski's Inequality

Let $f(x)$ and $g(x)$ be L^2 on the interval *I*, then

$$\left(\int_I |f(x)+g(x)|^2\,dx\right)^{1/2} \le \left(\int_I |f(x)|^2\,dx\right)^{1/2} + \left(\int_I |g(x)|^2\,dx\right)^{1/2}$$

Theorem 6: If $f(x)$ is continuous on a closed, finite interval *I*, then $f(x)$ is uniformly continuous on *I*, and its bounded on *I*; that is there exist a number $M > 0$ such that $|f(x)| < M \;\; \forall x \in I$.

جامعة النجاح الوطنية
العليا كلية الدراسات

النظرية الرياضية للمويجات

إعداد
بثينة محمد حسين غنام

إشراف
د. أنور صالح

قدمتْ هذه الأطروحة استكمالا لمتطلبات درجة الماجستير في الرياضيات بكليـــة الدراســات العليا في جامعة النجاح الوطنية في نابلس. فلسطين.

2009

النظرية الرياضية للمويجات

إعداد

بثينة محمد حسين غنام

إشراف

د. أنور صالح

الملخص

المويجات هي عبارة عن اقترانات تحقق شروط معينة وتستخدم في تمثيل ومعالجة الاقترانات والإشارات إضافة إلى ضغط البيانات والصور كما في تخصصات مثل: الرياضيات والفيزياء والحاسوب والهندسة والطب. وقد كان الحافز لدراسة تحويلات المويجات هو محاولة التغلب على بعض نقاط الضعف في تحويلات فوربيير التقليدية لتمثيل الاقترانات أو الإشارات مثل سرعة التقارب وظاهرة Gibbs. إضافة لذلك، أظهرت تحويلات المويجات تفوق على تحويلات فوربيير. ففي كثير من التطبيقات أثبت من الناحية النظرية والعملية أن سرعة تقارب تحويلات المويجات اكبر من سرعة تقارب تحويلات فوربيير ما يؤدي إلى معالجة أدق وأسرع للإشارات أو البيانات. في هذه الأطروحة نراجع الأسس الرياضية لتحويلات المويجات وتحويلات فوربيير. ونعمل مقارنة نظرية بين كلا التحويلين مثبتين تفوق تحولات المويجات على تحويلات فوربيير من حيث الدقة وسرعة التقارب في كثير من التطبيقات.