Order Number 9432355

Sampling theory in wavelet subspaces

Luo, Hua, Ph.D.

City University of New York, 1994

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

SAMPLING THEORY IN WAVELET SUBSPACES

by

HUA LUO

A dissertation submitted to the Graduate Faculty in Computer Science in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The City University of New York

1994

This manuscript has been read and accepted for the Graduate Faculty in Computer Science in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy.

Apr. 20, 94	
Date	Chair of Examining Committee
April 20, 1994	
Date	Executive Officer

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

SAMPLING THEORY IN WAVELET SUBSPACES

by

Hua Luo

Advisor: Professor Charles R. Giardina

In this dissertation, a method for creating sampling theorems in wavelet subspaces is given. Several sufficient conditions are provided for the sampling theorems to hold. The notion of continuous multiresolution analysis is introduced and the associated sampling theorems are extended to different (scale) wavelet subspaces. Aliasing error analysis techniques based on the idea of "band covering" are developed.

A useful and easy to apply sampling theorem is proven in the wavelet subspaces created by a special class of scaling functions, namely bandlimited sampling scaling functions. Another important sampling theorem is developed for scaling functions with raised cosine spectrum. The connection between the sampling theory and the problem of intersymbol interference is presented.

This Dissertation is Dedicated

to

My Wife Qing

and

Our First Child

ACKNOWLEDGMENTS

My deepest gratitude is extended to my mentor, Dr. Charles R. Giardina for his vision, support, encouragement and help in completing this study.

A grateful acknowledgement is expressed to my supervisory committee members: Dr. Charles R. Giardina, Dr. Stanley Habib, Dr. Michael Kress, Dr. Charles Suffel for their professional advice, generous help and support.

A special thank is extended to Dr. Albert Blank for proofreading the dissertation.

A bondless appreciation is extended to my uncle and aunt, Richard and Flora Chen, who made it possible for me to study in America.

Finally, my indebtness to my wife Qing and my parents Shaoji Luo and Yuehua Chen, whose love and support enabled me to persevere and see this project through to completion.

TABLE OF CONTENTS

LIST OF SYMBOLS

R	set of real numbers
R^+	set of positive real numbers
Z	set of integers
$[t]$	largest integer no bigger than t
$\int_E$	integral over E
$\bigvee_E f$	total variation of f on set E
$f(t^+)$	right limit of f at t
$f(t^-)$	left limit of f at t
$\{f(n)\}$	sequence of numbers
$\{f_n(t)\}$	sequence of functions
Σ_n or $\sum_n$	means $\sum_{n=-\infty}^{\infty} = \lim_{N\to\infty}\sum_{n=-N}^{N}$ in all cases excluding discrete convolution
$\hat{f}$, $\mathcal{F}f$	Fourier transform of function f
$\mathcal{F}^{-1}f$	inverse Fourier transform of function f
$\hat{f}^*$	discrete Fourier transform of $\{f(n)\}$, that is $\hat{f}^*(\omega)=\sum_n f(n)e^{-i\omega n}=\lim_{N\to\infty}\sum_{n=-N}^{N} f(n)e^{-i\omega n}$.
$f^{(N)}$	N-th derivative of function f

$C^N(R)$	collection of all functions $f^{(N)}$ is continuous
$L^p(E)$	collection of all measurable functions such that $\int_E \lvert f(t)\rvert^p dt<\infty$
L^2	means $L^2(R)$
$\lVert f\rVert_p$	p-norm of functions in $L^p(E)$ $$\lVert f\rVert_p=\begin{cases}\{\int_E \lvert f(x)\rvert^p dx\}^{1/p} & \text{for } 1\le p<\infty\\ \operatorname{ess\,sup}_{x\in E}\lvert f(x)\rvert & \text{for } p=\infty\end{cases}$$
(f,g)	inner product
l^p	infinite sequences $c=\{c_1,c_2,c_3,...\}$ such that $\lVert c\rVert_p= (\sum_{n=1}^{\infty}\lvert c_n\rvert^p)^{1/p}<\infty$
l^∞	space of all infinite sequences which are bounded, that is $\lVert c\rVert_\infty=\sup_n\{\lvert c_n\rvert\}<\infty$
ϕ	scaling function
ψ	wavelet
$\psi_{j,k}(t)$	$2^{j/2}\psi(2^j t-k)$
$\psi^{j,k}$	bi-orthogonal sequence of $\psi_{j,k}$
$\tilde{\psi}$	dual wavelet of ψ
$P(z)$	two-scale symbol of the scaling function

$Q(z)$	two-scale symbol of the wavelet
$\{V_j\}$	sequence of spaces generated by a (discrete) MRA
$\{U_\alpha\}$	index of spaces generated by a (continuous) MRA
$\{W_j\}$	sequence of spaces where the wavelets live
$x \in A$	x is an element of A
$A \subset B$	A is a subset of B
$x \notin A$	x is not an element of A
$A \not\subset B$	A is not a subset of B
$A \Leftrightarrow B$	A if and only if B
▌	Q.E.D.
$f_n \rightarrow f$	f_n converges to f
$a \neq b$	a is not equal to b
$A \cap B$	intersection
$A \cup B$	union
$A \dotplus B$	direct sum
$Clos_{L^2(R)}(A)$	$L^2(R)$ closure of A
$[a,b]$	close interval
(a,b)	open interval
supp f	a subset A of R such that $f(t)=0$ on the complement of A

$f(t)=O(t^p)$	$\lvert f(t)/t^p \rvert < M$ for some $M>0$ as $\lvert t \rvert \to \infty$
PW_π	Paley-Wiener space of functions bandlimited to $[-\pi,\pi]$
$[x_n]$	closed linear span of a vector sequence $\{x_n\}$
X^*	dual space of X (the space of all bounded linear functionals on X)
δ_{mn}	equals to 1 if $m=n$ and equal to 0 if $m\neq n$

LIST OF FIGURES

1. Introduction

The theory of wavelets emerged through the years, especially in the past 7-8 years from research in mathematics, physical science and signal processing. This theory attracts scientists in many disciplines. Some view wavelets as another way of representing functions, others use it as a tool for time-frequency or time-scale analysis. The subject is rapidly developing and has significant application value.

The mathematical ideas and techniques used in wavelet analysis can be traced back to the 1930s from the results of Haar, Hardy, Littlewood and Paley. Early on, these ideas didn't converge on a congruent theory and went unnoticed except to the specialist. Not until about 50 years later Yves Meyer put wavelets into a general theory and Stephane Mallat found the potential application of Multiresolution analysis. Occurrences such as this are not rare in the history of science. It compares to the formulation of the theory of distributions by Schwartz, where in 1951, he formalized the use of delta functions in engineering and physics. It also compares to the recovery of Whittaker's Sampling theorem by C. Shannon forming the field of information theory and to the recovery of the work of Hausdorff, Pierre Fatou and Gaston Julia by Mandelbrot forming the theory of Fractals.

The theory of wavelets as a part of Harmonic analysis has a long rich history. It started when Joseph Fourier [1807] claimed that every 2π-period function can be written in the form

$$f(t)=a_0+\sum_{k=1}^{\infty}(a_k\cos kt+b_k\sin kt).$$

Or formally, suppose f(t) is a 2π-period function, define the (Fourier) coefficients a_0, a_k, b_k $(k>0)$ by

$$a_0=\frac{1}{2\pi}\int_0^{2\pi}f(t)\,dt,$$

$$a_k=\frac{1}{\pi}\int_0^{2\pi}f(t)\cos kt\,dt, \qquad b_k=\frac{1}{\pi}\int_0^{2\pi}f(t)\sin kt\,dt,$$

then the partial sum

$$[S_n(f)](t)=a_0+\sum_{k=1}^{n}(a_k\cos kt+b_k\sin kt)$$

converge to f(t) pointwise and is denoted by $S_n(t)\to f$ (actually mathematicians of the eighteen century: Bernoulli, Euler, Lagrange and others, knew 'experimentally' that for some 'simple' functions, the above claim is true). Fourier also showed how the series could be used to solve some linear partial differential equations.

When Fourier announced this surprising result, neither the notion of function nor that of integral had a precise definition. His result helped the formulation of these concepts.

The first rigorous proof of convergence under rather general conditions is due to Dirichlet [1837]. The following is a consequence.

Dirichlet's result. If f is continuous and has a bounded derivative, then

$$[S_n(f)](t) = a_0 + \sum_{k=1}^{n} (a_k \cos kt + b_k \sin kt) \to f(t) \quad \text{as } n \to \infty$$

at all points (p3, Körner [1988]).

However, it turned out that the conditions on f cannot be relaxed altogether for in 1873 Paul Du Bois-Raymond constructed the following counter example.

Du Bois-Raymond's example. There exists a continuous function f such that the Fourier series of f diverging at the origin, here

$$\lim_{n \to \infty} \sup\, [S_n(f)](0) = \infty$$

(p3-4, Körner [1988]).

So, the Fourier series does not convergence in the naive sense of Fourier. Numerous results such as these have been produced (p74, Körner [1988]). The most striking (and difficult) result in the field of Fourier analysis is L. Carleson's proof of Lusin's conjecture. Lusin conjectured that the Fourier series of L^2 function converge almost everywhere (a.e.) [1913] and Carleson proved this fact [1966] (p20, Hunt [1976]). The development of Fourier theories and associated techniques greatly influenced the advancement of mathematics, physics, engineering and also affected all aspect of human lives.

Here, sampling theory in wavelet subspaces is presented. Chapter 2, section 2.1 starts with the evolution of wavelet analysis from Fourier analysis. Then a tool for finding wavelets, the (discrete) multiresolution analysis is given in section 2.2. Some examples related to the problem of sampling theory is presented in section 2.3.

Chapter 3 is the main part of the dissertation. In section 3.1, the "proper choice condition" problem is presented. Here, theorems (3.1, 3.2 and several corollaries) are proved providing sufficient conditions for which the sampling theorems holds. In order to deal with the problem of continuous scaling, the notion of a (continuous) multiresolution analysis is introduced in section 3.2 by the author and sampling theorems are extended to different (scale) wavelets subspaces (theorem 3.3 and 3.4). In section 3.3, the aliasing error analysis techniques based on "band covering" are developed. Examples using the theory developed in 3.1-3.3 are given in section 3.4.

In chapter 4 the special class of bandlimited sampling scaling functions are used in sampling theory. In section 4.1, additional theorems are proven based on the theorems in chapter 3. In section 4.2, the scaling function with raised cosine spectrum is introduced as a special example of bandlimited sampling scaling function. In section 4.3, the connection between the sampling theory and the problem of intersymbol interference is presented.

Appendix A and B are included to make the dissertation self-contained. Here the mathematical theory needed to understand the proofs in the dissertation is given along with extensive references for further reading.

2 Wavelet and Multiresolution Analysis

2.1. From Fourier Analysis to Wavelet Analysis

Consider the function space $L^2[-\pi,\pi]$, that is the class of all measurable functions f such that

$$\int_{-\pi}^{\pi}|f(x)|^2dx<\infty.$$

Define

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

and

$$S_N(x)=\sum_{n=-N}^{N}c_n e^{ixn},$$

then the function S_N converges to f in the sense of $L^2[-\pi,\pi]$, that is

$$\lim_{N\to\infty}\int_{-\pi}^{\pi}|f(x)-S_N(x)|^2dx=0.$$

For a detailed discussion about the space $L^2[-\pi,\pi]$ and Fourier series, see appendix A and B.3.

Hence any $f\in L^2[-\pi,\pi]$ can be written as

$$f(x)=\sum_{n=-\infty}^{\infty}c_n e^{ixn}$$

in the sense of $L^2[-\pi,\pi]$ for some complex sequence $\{c_n\}$. Use the terminology in appendix A, e^{ixn} is called a basis of $L^2[-\pi,\pi]$. Since e^{ixn} is just a dilation of the function e^{ix}, it can be said e^{ix} "generates" the space $L^2[-\pi,\pi]$.

Now consider the space $L^2(R)$, which is defined as the class of all measurable functions such that

$$\int_{-\infty}^{\infty} |f(x)|^2 dx < \infty .$$

Like in the space $L^2[-\pi,\pi]$, where a single function e^{ix} "generates" the whole space, some kind of function, or "wave" is desired which "generates" all of $L^2(R)$. Since every function $f(x)$ in $L^2(R)$ must "decay" to zero in energy as x goes to infinity, it intuitively follows that some kind of small wave, or "wavelet" would generate all of $L^2(R)$.

Among the simplest ways that a function ψ can "generate" all of $L^2(R)$ is that ψ be a so called R-function. Here the collection of functions

$$2^{j/2}\psi(2^j x-k), \quad j,k\in Z$$

forms a basis of $L^2(R)$.

Definition 2.1. A function $\psi \in L^2(R)$ is called an **R-function,** if the associated functions $\{\psi_{j,k}\}$,

$$\psi_{j,k}(x)=2^{j/2}\psi(2^j x-k), \quad j,k\in Z$$

form a Riesz basis of $L^2(R)$ (For a detailed discussion about Riesz basis, see appendix A.8.).

Since $\{\psi_{j,k}\}$ is a basis of $L^2(R)$, its biorthogonal sequence $\{\psi^{j,k}\}$ is also a basis of $L^2(R)$ (called the dual basis) and every $f \in L^2(R)$ has the following (unique) series expansions

$$f(x) = \sum_{j,k=-\infty}^{\infty} \langle f, \psi_{j,k} \rangle \psi^{j,k}(x) .$$

For the function ψ to be qualified as a wavelet, there must exist some function $\tilde{\psi} \in L^2(R)$, such that $\{\psi^{j,k}\}$ is obtained from $\tilde{\psi}$ by

$$\psi^{j,k}(x) = \tilde{\psi}_{j,k}(x)$$

where, as usual, the notation

$$\tilde{\psi}_{j,k}(x) = 2^{j/2} \tilde{\psi}(2^j x - k)$$

is used. But in general, $\tilde{\psi}$ may not exist (p13, Chui [1992a]).

The following is the precise definition of "wavelet".

Definition 2.2. An R-function ψ is called an **R-wavelet (or wavelet)**, if there exists a function $\tilde{\psi} \in L^2(R)$, such that $\{\psi_{j,k}\}$ and $\{\tilde{\psi}_{j,k}\}$, as defined above, are dual bases of $L^2(R)$.

If ψ is a R-wavelet, then $\tilde{\psi}$ is called a **dual wavelet** corresponding to ψ. It is clear that the dual wavelet $\tilde{\psi}$ is unique and is itself an R-wavelet.

The most important basis of a space is an o.n. basis (appendix A). If $\{\psi_{j,k}\}$ is an o.n. basis of $L^2(R)$, then $\psi^{j,k}=\psi_{j,k}$, so $\tilde{\psi}$ exists and $\tilde{\psi}=\psi$. Hence if $\{\psi_{j,k}\}$ is an o.n. basis, ψ is always a wavelet, and in fact, an orthonormal one.

Definition 2.3. If $\{\psi_{j,k}\}$ is an o.n. basis of $L^2(R)$, then ψ is called an **orthonormal wavelet.**

The next simplest family of wavelets are the semi-orthogonal wavelets.

Definition 2.4. A wavelet ψ is called a **semi-orthogonal (s.o.) wavelet** if the Riesz basis $\{\psi_{j,k}\}$ satisfies

$$\langle \psi_{j,k}, \psi_{l,m} \rangle = 0, \quad j \neq l; \qquad j,k,l,m \in Z.$$

Using the following orthonormalization procedure, an orthonormal wavelet can be generated from a semi-orthogonal wavelet.

Suppose ψ is an s.o. wavelet and if $\psi^{\perp}$ is defined from its Fourier transform

$$\hat{\psi}^{\perp}(\omega) = \frac{\hat{\psi}(\omega)}{\sqrt{\sum_k |\hat{\psi}(\omega+2\pi k)|^2}}$$

then $\psi^{\perp}$ is an o.n. wavelet. (p79-80, Chui [1992a])

The first example of orthonormal wavelet was given by Haar [1909], where ψ is chosen to be

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

For a long time, there were no other o.n. wavelets available. Moreover there was no evidence that other wavelets might exist. This lasted until the discovery of the Franklin wavelets by Strömberg [1980].

Most wavelets today are found using the so called "multiresolution analysis" technique, initiated by Yves Meyer [1986] and applied in image processing by Stephane Mallet [1988]. Not every wavelet can be found by the multiresolution analysis. However all wavelets in this dissertation are generated from a multiresolution analysis.

2.2. Discrete Multiresolution Analysis

Given the wavelet $\psi \in L^2(R)$, for any fixed j, define W_j to be the $L^2(R)$ closure of the linear span of the $\{\psi_{j,k}\}$, that is

$$W_j = [\psi_{j,k} : k \in Z]$$

Then this family of subspaces form a direct sum decomposition of $L^2(R)$,

$$L^2(R) = \cdots \dotplus W_{-1} \dotplus W_0 \dotplus W_1 \dotplus \cdots$$

where $\dotplus$ is the direct sum operator of vector spaces. Thus for any $f \in L^2(R)$

$$f(x) = \cdots + g_{-1}(x) + g_0(x) + g_1(x) + \cdots$$

where $g_j \in W_j$ and uniquely determined.

Define

$$V_j = \cdots \dotplus W_{j-3} \dotplus W_{j-2} \dotplus W_{j-1}$$

Then the spaces V_j satisfy the following theorem.

Theorem 2.1. The family of subspaces V_j have the following properties:

(1) V_j are closed and nested i.e. $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$;

(2) $\mathrm{Clos}_{L^2}(\bigcup_{j \in Z} V_j) = L^2(R)$;

(3) $\bigcap V_{j \in Z} = \{0\}$;

(4) $V_{j+1} = V_j \dotplus W_j$;

(5) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$, $j \in Z$.

(see p120, Chui[1992a]).

The properties in theorem 2.1. are necessary conditions for the existence of a wavelet ψ. In order to better understand properties of the spaces $\{V_j\}$, the so called "scaling function" ϕ is introduced. Formally, a function $\phi \epsilon V_0$ is to be found such that

$$\{\phi(\cdot - k) : k \epsilon Z\}$$

is a Riesz basis of V_c. This leads to the idea of multiresolution analysis.

A sequence of subspaces $\{V_j\}$ of $L^2(R)$ will be introduced. This is with the understanding that they are no longer defined using the subspaces W_j.

Definition 2.5. The sequence $\{V_j\}$ is said to form a **(discrete) multiresolution analysis (MRA)** of $L^2(R)$ if the following five conditions are satisfied.

(1°) V_j are closed and nested i.e. $\cdots \subset V_{-1} \subset V_0 \subset V_1 \subset \cdots$;

(2°) $\mathrm{Clos}_{L^2}(\bigcup_{j \epsilon Z} V_j) = L^2(R)$;

(3°) $\bigcap V_{j \epsilon Z} = \{0\}$;

(4°) $f(x) \epsilon V_j \Leftrightarrow f(2x) \epsilon V_{j+1}$, $j \epsilon Z$;

(5°) There is a function ϕ, such that $\{\phi(x-k)\}$ is a Riesz basis of V_0 with Riesz bounds A and B.

Here ϕ is called a **scaling function,** and it generates a (discrete) multiresolution analysis (MRA) V_j, $j\epsilon Z$. If, in addition, A=B=1, then $\{\phi(x-k)\}$ becomes an orthonormal (o.n.) basis of V_0, and ϕ is called an **orthonormal scaling function.**

Notice by setting

$$\phi_{j,k}(x)=2^{j/2}\phi(2^j x-k),$$

it follows from (4°) that for each fixed $j\epsilon Z$, the family

$$\{\phi_{j,k}:k\epsilon Z\}$$

is also a Riesz basis of V_j with the same Riesz bounds A and B. As a consequence, the MRA also satisfies the following property:

(6°) $f(x)\epsilon V_j \Leftrightarrow f(x+1/2^j)\epsilon V_j$, $j\epsilon Z$.

Let ϕ generate a MRA. Since $\phi\epsilon V_0 \subset V_1$ and $\{2^{\frac{1}{2}}\phi(2x-k)\}$ is a Riesz basis of V_1, there exists a unique l^2 sequence $\{p_k\}$ such the **Two scale relation involving the scaling function** ϕ holds:

$$\phi(x)=\sum_{k=-\infty}^{\infty} p_k\phi(2x-k).$$

This sequence $\{p_k\}$ is called the **two scale sequence for ϕ** and the corresponding **two scale symbol P of ϕ** is defined as

$$P(z)=P_\phi(z)=\frac{1}{2}\sum_{k=-\infty}^{\infty} p_k z^k.$$

This is similar to the z-transform or the moment factorial generating function of a sequence except that the normalization constant 1/2 is used to simplify the following Fourier transform formulation.

$$\hat{\phi}(\omega) = P(e^{-i\omega/2})\hat{\phi}(\omega/2).$$

Thus P and ϕ uniquely determined each other as well as the discrete MRA they generate.

As previously mentioned, if a wavelet exists, it lives in the spaces W_j, where $V_{j+1}=V_j+W_j$. If a discrete MRA is chosen properly, then a wavelet ψ can be determined.

If the wavelet ψ is found, since $\psi \in W_0 \subset V_1$, there exists another unique l^2 sequence $\{q_k\}$ such that the **Two scale relation of the wavelet** ψ holds:

$$\psi(x) = \sum_{k=-\infty}^{\infty} q_k \phi(2x-k).$$

Similarly $\{q_k\}$ is called the **two scale sequence of** ψ and the **two scale symbol of** ψ is defined as

$$Q(z) = Q_\psi(z) = \frac{1}{2}\sum_{k=-\infty}^{\infty} q_k z^k.$$

When ϕ is an o.n. scaling function, the relationship between $\{p_k\}$ and $\{q_k\}$ is very simple, here

$$q_n = (-1)^n \overline{p_{1-n}}.$$

Since both $\phi(2x)$ and $\phi(2x-1)$ are in V_1 and $V_1 = V_0 \dotplus W_0$, there exists four l^2 sequences $\{a_{-2k}\}, \{b_{-2k}\}, \{a_{1-2k}\}$ and $\{b_{1-2k}\}$ such that

$$\phi(2x) = \sum_k [a_{-2k}\phi(x-k) + b_{-2k}\psi(x-k)]$$

$$\phi(2x-1) = \sum_k [a_{1-2k}\phi(x-k) + b_{1-2k}\psi(x-k)]$$

for all $x \in R$. The above two formulas can be combined into a single **decomposition formula:**

$$\phi(2x-l) = \sum_k [a_{l-2k}\phi(x-k) + b_{l-2k}\psi(x-k)], \quad l \in Z$$

This is equivalent to

$$\phi(2^{j+1}x-l) = \sum_k [a_{l-2k}\phi(2^j x-k) + b_{l-2k}\psi(2^j x-k)], \quad j, l \in Z$$

Two pairs of sequences $(\{p_k\},\{q_k\})$ and $(\{a_k\},\{b_k\})$ correspond to the two scale relationships and the decomposition formula are both uniquely determined by the direct sum relationship $V_{j+1} = V_j \dotplus W_j$. These sequences are used to formulate the following reconstruction and decomposition algorithms.

Any $f_j \in V_j$ and $g_j \in W_j$ have the unique representations

$$f_j(x) = \sum_k c_k^j \phi(2^j x - k)$$

$$g_j(x) = \sum_k d_k^j \psi(2^j x - k)$$

Here the normalization coefficient $2^{j/2}$ is dropped to facilitate the computation. Substitute the decomposition formula into

$$f_j(x) = \sum_l c_l^j \phi(2^j x - l)$$

gives

$$f_j(x) = \sum_l c_l^j \sum_k [a_{l-2k} \phi(2^{j-1}x - k) + b_{l-2k} \psi(2^{j-1}x - k)]$$

$$= \sum_k \sum_l c_l^j [a_{l-2k} \phi(2^{j-1}x - k) + b_{l-2k} \psi(2^{j-1}x - k)]$$

This leads to the **decomposition algorithm:**

$$\begin{cases} c_k^{j-1} = \sum_l a_{l-2k} c_l^j \\ d_k^{j-1} = \sum_l b_{l-2k} c_l^j \end{cases}$$

Using the symbols $c^j = \{c_k^j\}$ and $d^j = \{d_k^j\}$, the decomposition algorithm can be visualized as:

$$\begin{array}{ccccccccc} & & d^{N-1} & & d^{N-2} & & & & d^{N-M} \\ & \nearrow & & \nearrow & & \nearrow & & \nearrow & \\ c^N & \rightarrow & c^{N-1} & \rightarrow & c^{N-2} & \rightarrow & \cdots & \rightarrow & c^{N-M} \end{array}$$

Fig 2.1. Wavelet decomposition algorithm

By using the two-scale relations

$$f_{j-1}(x)=\sum_{l} c_l^{j-1}\phi(2^{j-1}x-l)$$

$$=\sum_{l} c_l^{j-1}\sum_{k} p_k\phi(2^j x-2l-k)$$

$$=\sum_{l} c_l^{j-1}\sum_{k} p_{k-2l}\phi(2^j x-k)$$

$$=\sum_{k}\sum_{l} c_l^{j-1}p_{k-2l}\phi(2^j x-k)$$

Similarly

$$g_{j-1}(x)=\sum_{k}\sum_{l} d_l^{j-1}q_{k-2l}\phi(2^j x-k)$$

If $f_j=f_{j-1}+g_{j-1}$, then **reconstruction algorithm** is obtained

$$c_k^j=\sum_{l}[p_{k-2l}c_l^{j-1}+q_{k-2l}d_l^{j-1}]$$

$$\begin{array}{ccccccc} d^{N-M} & & d^{N-M+1} & & & d^{N-1} & \\ & \searrow & & \searrow & & & \searrow \\ c^{N-M} & \rightarrow & c^{N-M+1} & \rightarrow & \cdots & c^{N-1} & \rightarrow c^{N} \end{array}$$

Fig 2.2. Wavelet reconstruction algorithm

2.3. Examples of Wavelets.

Example 2.1. Haar wavelet.

The o.n. scaling function is $\phi(x)=\chi_{[0,1)}$. Space V_0 consists of all the step functions with the step width being 1 and the wavelet is

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

EXAMPLE 2.2. Cardinal series.

The o.n. scaling function in this case is

$$\phi(t)=\frac{\sin\pi t}{\pi t}$$

and V_0 is a Paley-Wiener space with functions bandlimited to $[-\pi,\pi]$. The wavelet is

$$\psi(t)=\frac{\sin\pi(t-1/2)-\sin 2\pi(t-1/2)}{\pi(t-1/2)}.$$

EXAMPLE 2.3. Franklin wavelets.

First form a multiresolution analysis $\{V_m\}$ by choosing the scaling function to be the second order B-spline

$\eta(t)=(1-|t-1|)\chi_{[0,2]}(t)$. The space V_0 consists of continuous piecewise affine functions with breaks at the integers. The Fourier transform of η is

$$\hat{\eta}(\omega)=\frac{\sin^2\omega/2}{(\omega/2)^2}e^{-i\omega}$$

and the scaling function for the Franklin wavelets is obtained from the orthonormalization procedure similar to the one used in deriving the o.n. wavelet from the s.o. wavelet. That is, ϕ is defined by employing its Fourier transform

$$\hat{\phi}(\omega)=\frac{\hat{\eta}(\omega)}{\sqrt{\sum_k|\hat{\eta}(\omega+2\pi k)|^2}}=\frac{\sin^2\omega/2}{(\omega/2)^2}\left(1-\frac{2}{3}\sin^2\frac{\omega}{2}\right)^{-\frac{1}{2}}.$$

EXAMPLE 2.4. Quadratic spline.

Haar and Franklin wavelets correspond to first and second order basic splines respectively. For the quadratic spline or basic third order spline the scaling function is η, and its Fourier transform

$$\hat{\eta}(\omega)=\left[\frac{1-e^{-i\omega}}{i\omega}\right]^3$$

η is not an o.n. scaling function, but it can be orthonormalized by the same process as used in example 2.3.

EXAMPLE 2.5. Lemarié-Meyer Wavelet.

Let $0<\varepsilon\leq\pi/3$, $0<A<1<B<\infty$, and N be an arbitrary positive integer. Choose any $\hat{\eta}\in C^N(R)$ where

(a) supp $\hat{\eta}=[-\pi-\varepsilon,\pi+\varepsilon]$;

(b) $\hat{\eta}(\omega)=1$ for $|\omega|\leq\pi-\varepsilon$; and

(c) $A\leq\sum_k|\hat{\eta}(\omega+2\pi k)|^2\leq B$

If ϕ whose Fourier transform is given by

$$\hat{\phi}(\omega)=\frac{\hat{\eta}(\omega)}{(\sum_k|\hat{\eta}(\omega+2\pi k)|^2)^{1/2}}$$

Then ϕ is an o.n. scaling function (p217, Chui [1992a]).

EXAMPLE 2.6. Bandlimited sampling scaling function

Choose any $\hat{\eta}\in C^N(R)$, N being any positive integer and

(a) supp $\hat{\eta}=[-\pi-\varepsilon,\pi+\varepsilon]$, $0<\varepsilon\leq\pi$;

(b) $\hat{\eta}(\omega)=1$ for $|\omega|\leq\pi-\varepsilon$;

(c) $A\leq\sum_k|\hat{\eta}(\omega+2\pi k)|^2\leq B$, $0<A<1<B<\infty$;

(d) $\hat{\eta}(\omega)\geq 0$; and

(e) $\hat{\eta}(\omega)>0$, for $|\omega|<(\pi+\varepsilon)/2$

If ϕ whose Fourier transform is given by

$$\hat{\phi}(\omega) = \frac{\hat{\eta}(\omega)}{(\sum_k |\hat{\eta}(\omega+2\pi k)|^2)^{1/2}}$$

then ϕ is an o.n. scaling function. Notice the construction of ϕ is very similar to Lemarié-Meyer scaling function. Actually, it is obtained by loosening the condition of ε, but adding conditions (d) and (e). The reason these changes are made is that this family of scaling functions will be extremely useful to the application of sampling theory which is developed in chapter 3. Chapter 4 is devoted to the analysis of the effects of the sampling theorems on this type of scaling function.

3 Sampling Theory and Error Analysis in Wavelet Subspaces

3.1. Sampling Theorems in V_0

In 1948-1949, Claude Shannon published several papers establishing the foundation of modern communication theory. The starting point of his theory is the following theorem:

If a function f(t) contains no frequencies higher than W cps, it is completely determined by giving its ordinates at a sequence of points spaced 1/2W second apart.

Though Shannon pointed out in those papers the equation

$$f(t)=\sum_{n=-\infty}^{\infty} f\left(\frac{n}{2W}\right)\frac{\sin\pi(2Wt-n)}{\pi(2Wt-n)}$$

had been proved at least 30 years earlier by Oxford mathematician J. Whittaker [1915]. The above equation is best known as Shannon sampling theorem. The reason people almost forget Whittaker totally is apparently because Shannon did such a good job in connecting the above theorem to the problem of communication.

Many extensions of this classical sampling theorem have been proposed. The possibility of extension to the wavelet subspace was first recognized by Gilbert Walter [1992].

Walter's result. If ϕ is a real continuous function such that

(i) $\phi(t)=O(|t|^{-1-\varepsilon})$ as $t\to\pm\infty$, where ε is some small positive number.

(ii) $\hat{\phi}^{*}(\omega)=\sum_n \phi(n)e^{-i\omega n}\neq 0$, for all $\omega\in R$.

(iii) ϕ is an o.n. scaling function generating the multiresolution system $\{V_m\}$, $m\in Z$. Then for $f\in V_0$,

$$f(t)=\sum_{n=-\infty}^{+\infty} f(n)S(t-n)$$

where

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^{*}(\omega)$$

and the convergence is uniform, i.e.

$$\lim_{N\to\infty}\|f(t)-\sum_{n=-N}^{N} f(n)S(t-n)\|_{\infty}=0$$

Walter noticed that V_0 is a reproducing kernel Hilbert space. He chose $\{S(t-n)\}$ to be the dual basis of $\{q(t,n)\}$, where $q(t,s)$ is the reproducing kernel of V_0, and then "proved" the above result.

The theorem as given above doesn't hold, since if

$$f(t)=\begin{cases} S(t) & t\neq 0 \\ 0 & t=0 \end{cases}$$

Then because $S\epsilon V_0$, using the above theorem gives

$$S(t)=\sum_n S(n)\,S(t-n)\,.$$

Since $\{S(t-n)\}$ is a basis of V_0, $S(n)=\delta_{0n}$. Using the above theorem again

$$f(t)=\sum_n f(n)\,S(t-n)=0\,.$$

which is impossible.

The reason the above counter example holds is that every function in L^2 belongs to an equivalence class. The proper selection of the representing function is critical for the sampling theory to be valid. This issue, herein called the "proper choice condition" problem, will be discussed in detail later. Nonetheless, Walter's result holds for most real life signals (see Corollary 2 of theorem 3.2.) and it provides great insight in what the sampling theorems in wavelet subspaces will look like.

Before the sampling theorems along with rigorous proofs are introduced, a formal presentation of the theorem will be given.

Assume that translates of a function ϕ, that is $\{\phi(x-n)\}$ form a basis for V_0. Then for any f in V_0, it is true that

$$f(t)=\sum_n a_n \phi(t-n) .$$

The objective is to show that

$$f(t)=\sum_n f(n) S(t-n)$$

where S(t) has Fourier transform

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega) .$$

The Discrete Fourier transform (DFT) $\hat{\phi}^*(w)$ of $\{\phi(n)\}$ is given by

$$\hat{\phi}^*(\omega)=\sum_n \phi(n) e^{-i\omega n}.$$

Instead of showing that

$$f(t)=\sum_n f(n) S(t-n) ,$$

the Fourier transform of this equation can be shown to hold:

$$\hat{f}(\omega)=\sum_n f(n) e^{-i\omega n}\hat{S}(\omega) .$$

The last equation can be written as

$$\hat{f}(\omega)=\hat{f}^*(\omega)\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$$

or

$$\hat{f}(\omega)\hat{\phi}^*(\omega)=\hat{f}^*(\omega)\hat{\phi}(\omega) .$$

Since

$$f(t)=\sum_k a_k\phi(t-k) ,$$

let

$$f(n)=\sum_k a_k\phi(n-k)$$

and

$$\hat{f}(\omega)=\sum_n a_n e^{-i\omega n}\cdot\hat{\phi}(\omega)$$

since the latter equation is the Fourier transform of f(t). Substitution of these equations into

$$\hat{f}(\omega)\hat{\phi}^*(\omega)=\hat{f}^*(\omega)\hat{\phi}(\omega)$$

gives

$$\sum_n a_n e^{-i\omega n}\sum_n \phi(n)\, e^{-i\omega n}\hat{\phi}(\omega)=\sum_n\sum_k a_k\phi(n-k)\, e^{-i\omega n}\hat{\phi}(\omega)$$

or

$$\sum_n a_n e^{-i\omega n}\sum_n \phi(n)\, e^{-i\omega n}=\sum_n\sum_k a_k\phi(n-k)\, e^{-i\omega n} .$$

The left hand side of the final equation is the product of the DFT of a and ϕ, on the right is the DFT of the convolution of a and ϕ. The equality follows by interchanging the order of summation on the right side:

$$\sum_n \sum_k a_k \phi(n-k)\, e^{-i\omega n} = \sum_k \sum_n a_k \phi(n-k)\, e^{-i\omega n}$$

$$= \sum_k a_k e^{-i\omega k} \sum_n \phi(n-k)\, e^{-i\omega(n-k)}$$

$$= \sum_k a_k e^{-i\omega k} \sum_k \phi(k)\, e^{-i\omega k}.$$

This formal proof will be made rigorous and with additional notation and lemmas.

Recall that the symbol

$$\sum_n \textit{ means } \sum_{n=-\infty}^{\infty} = \lim_{N\to\infty} \sum_{n=-N}^{N}$$

thus

$\sum_n a_n = A$ *implies that for any* $\varepsilon > 0$ *there is a* $N > 0$ *s.t.* $\sum_{|n|>N} a_n < \varepsilon$.

LEMMA 0. If

$$\sum_n a_n = A \textit{ and } \sum_n b_n = B,$$

then

$$\sum_n (a_n+b_n)=A+B \text{ and } \sum_n ca_n=cA \text{ for any fixed } c.$$

(see p63, Rudin [1964]).

DEFINITION 3.1. The sequence of functions $\{f_n\}$, where $f_n \in L^2$ for all n, **converge to f in $L^2(E)$** is denoted by $f_n \to f$ in $L^2(E)$ and means that

$$\int_E |f_n(x)-f(x)|^2 dx \to 0 \text{ as } n \to \infty,$$

where E can be R or $[-\pi,\pi]$. By the completeness of $L^2(E)$, $f \in L^2(E)$. The convergence in $L^2(E)$ as defined above is sometimes called strong convergence in $L^2(E)$ or convergence in the norm of $L^2(E)$.

DEFINITION 3.2. The two functions f and g are **equal in $L^2(E)$**, denoted by $f=g$ in $L^2(E)$, if

$$\int_E |f(x)-g(x)|^2 dx = 0.$$

By the triangle inequality, it follows that if $f_n \to f$ in $L^2(E)$ and $f_n \to g$ in $L^2(E)$, then $f=g$ in $L^2(E)$.

LEMMA 1. If $f_n \to f$ in $L^2(E)$ and g is bounded on E, then $f_n g \to fg$ in $L^2(E)$.

Proof. Since g is bounded, that is $|g(x)|<M$ for some $M>0$ on E,

$$\int_E |f_n(x)\, g(x) - f(x)\, g(x)|^2 dx \le M^2 \int_E |f_n(x) - f(x)|^2 dx.$$

Hence $f_n g \to fg$ *in* $L^2(E)$ *whenever* $f_n \to f$ *in* $L^2(E)$. ∎

LEMMA 2. $f_n \to f$ in $L^2(R)$ if and only if $\hat{f}_n \to \hat{f}$ in $L^2(R)$.

Proof. By Parseval's Equality, $\|\hat{f}\|_2 = (2\pi)^{½}\|f\|_2$ for any $f \in L^2(R)$. Hence $\|\hat{f}-\hat{f}_n\|_2 = (2\pi)^{½}\|f-f_n\|_2$. Therefore $f_n \to f$ in $L^2(R)$ if and only if $\hat{f}_n \to \hat{f}$ in $L^2(R)$. ∎

DEFINITION 3.3. The symbol $\hat{f}^*(w)$ associated with the sequence $\{f(n)\}$ sometimes called the discrete Fourier transform of $\{f(n)\}$ is defined as

$$\hat{f}^*(\omega) = \sum_n f(n)\, e^{-i\omega n} = \lim_{N\to\infty} \sum_{n=-N}^{N} f(n)\, e^{-i\omega n}.$$

For $\{f(n)\} \in l^1$, $\hat{f}^*(w)$ is a continuous function well defined by the above equation (Lemma 3). Moreover when $\{f(n)\} \in l^2$, $\hat{f}^*(w)$ is equal to a function in $L^2[-\pi,\pi]$ (Riesz-Fisher Theorem). In both cases, $\hat{f}^*(w)$ is
2π-periodic.

LEMMA 2'. The sequence (with index N) involving the sequence $\{f_N(n)\}$ converge to $\{f(n)\}$ in l^2 if and only if $\hat{f}^*_N(w)$ converges to $\hat{f}^*(w)$ in $L^2[-\pi,\pi]$.

Proof. The result follows from Parseval equality involving Fourier series,

$$\sum_n |f(n)|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |\hat{f}^*(\omega)|^2 d\omega . \qquad \blacksquare$$

LEMMA 3. If $\{g(n)\}\in l^1$, then

$$\hat{g}^*(\omega) = \sum_n g(n)\, e^{-i\omega n}$$

is 2π-periodic, continuous and bounded. If, in addition $\hat{g}^*(w)\neq 0$, then there exist some $m>0$

$$\hat{g}^*(\omega) \geq m.$$

Proof. The periodicity is proved by simply noticing that

$$\hat{g}^*(\omega+2\pi) = \sum_n g(n)\, e^{-i(\omega+2\pi)n} = \sum_n g(n)\, e^{-i\omega n} = \hat{g}^*(\omega) .$$

So if the continuity of $\hat{g}^*$ can be proved, it readily follows that $\hat{g}^*$ is bounded. Considering

$$|\hat{g}^*(\omega+\delta) - \hat{g}^*(\omega)| = |\sum_n g(n)\, e^{-i(\omega+\delta)n} - \sum_n g(n)\, e^{-i\omega n}|,$$

by lemma 0, the above equation becomes

$$|\sum_n g(n)\, e^{-i\omega n}(e^{-i\delta n}-1)| \leq \sum_n |g(n)\,(e^{-i\delta n}-1)|.$$

Since $\{g(n)\}\in l^1$ and $|e^{-i\delta n}-1|\le 2$, for any $\varepsilon>0$, there exists $N>0$ such that

$$\sum_{|n|>N}|g(n)(e^{-i\delta n}-1)|<\varepsilon .$$

From the continuity of the sinusoid functions, there exists $\tau>0$, such that $\tau N<\pi$ and for $|\delta|<\tau$

$$|\sin\delta N/2|<\frac{\varepsilon}{\sum_{|n|\le N}|g(n)|} .$$

Since for $|n|\le N$

$$|\delta N|<\pi,\ hence\ |\sin\delta n/2|\le|\sin\delta N/2|$$

and because

$$|e^{-i\delta n}-1|=|e^{-i\delta n/2}(e^{-i\delta n/2}-e^{i\delta n/2})|=2|\sin\delta n/2|,$$

so for $|\delta|<\tau$

$$\sum_{|n|\le N}|g(n)(e^{-i\delta n}-1)|=\sum_{|n|\le N}|2g(n)\sin\delta n/2|\le\sum_{|n|\le N}2|g(n)||\sin\delta N/2|<2\varepsilon .$$

Hence for $|\delta|<\tau$

$$|\hat{g}^*(\omega+\delta)-\hat{g}^*(\omega)|\le\sum_{n}|g(n)(e^{-i\delta n}-1)|<\varepsilon+2\varepsilon=3\varepsilon .$$

So $\hat{g}^*$ is continuous.

Whenever a function h is continuous, the inequality

$$|h(t+\Delta t)|-|h(t)|| \le h(t+\Delta t)-h(t)\ |$$

shows that the function $|h|$ is also continuous. Since $\hat{g}^*(\omega)$ is periodic and continuous, $|\hat{g}^*(\omega)|$ is also periodic and continuous, Hence $|\hat{g}^*(\omega)|$ has a minimum. Therefore if in addition $\hat{g}^*(\omega)\neq 0$,

$$\hat{g}^*(\omega) \geq m$$

for some $m>0$. ∎

DEFINITION 3.4. The discrete convolution of f and g is formally defined by

$$h(n)=\sum_k f(k)\,g(n-k)=\lim_{\substack{M\to\infty\\ N\to\infty}}\sum_{k=-M}^{N} f(k)\,g(n-k)\,.$$

By a change of variables

$$h(n)=\lim_{\substack{M\to\infty\\ N\to\infty}}\sum_{k=n-N}^{n+M} f(n-k)\,g(k)=\lim_{\substack{M'\to\infty\\ N'\to\infty}}\sum_{k=-M'}^{N'} f(n-k)\,g(k)=\sum_k f(n-k)\,g(k)\,.$$

Hence whenever discrete convolution is involved, the symbol

$$\sum_k \text{ means } \lim_{\substack{M\to\infty\\ N\to\infty}}\sum_{k=-M}^{N}$$

and all the other cases $\sum_k$ means as previously mentioned

symmetric summation, i.e.

$$\sum_{k} \text{ means } \lim_{N\to\infty}\sum_{k=-N}^{N}$$

LEMMA 4. If $\{f(n)\}\in l^2$, $\{g(n)\}\in l^1$, then the discrete convolution of f and g

$$h(n)=\sum_{k} f(k)\,g(n-k)$$

exists and satisfies

$$\sum_{n} h^2(n) \le [\sum_{n} |g(n)|]^2 \sum_{k} f^2(k)\,.$$

Proof. First consider the case when $f(n)\ge 0$, $g(n)\ge 0$ for all n. If a series involving nonnegative terms converges, then any rearrangement also converge to the same sum.

$$\sum_{n} h^2(n) = \sum_{n} (\sum_{k} f(n-k)\,g(k))^2$$

$$=\sum_{n} [\sum_{i,j\in Z} f(n-i)\,f(n-j)\,g(i)\,g(j)]$$

$$=\sum_{i,j\in Z} [\sum_{n} f(n-i)\,f(n-j)\,g(i)\,g(j)]$$

$$=\sum_{i,j\in Z} [g(i)\,g(j)\sum_{n} f(n-i)\,f(n-j)]$$

By the CBS inequality,

$$\sum_{n} f(n-i)\,f(n-j) \le \sqrt{\sum_{n} f^2(n-i)}\cdot\sqrt{\sum_{n} f^2(n-j)} = \sum_{n} f^2(n)$$

for any i,j∈Z. Therefore

$$\sum_n h^2(n) \le [\sum_{i,j\in Z} g(i)g(j)]\sum_n f^2(n) = [\sum_n g(n)]^2 \sum_n f^2(n)$$

In the case when some of the f(n) or g(n) are negative, define

$$h_\$(n) = \sum_k |f(k)||g(n-k)|,$$

then

$$|h(n)| = |\sum_k f(k)g(n-k)| \le \sum_k |f(k)g(n-k)| = h_\$(n) \text{ and}$$

$$\sum_n h^2(n) \le \sum_n h_\$^2(n) \le [\sum_n |g(n)|]^2 \sum_n f^2(n).$$ ∎

REMARK. If $f,g\in L^2[-\pi,\pi]$, fg needn't be in $L^2[-\pi,\pi]$ ($fg\in L^1[-\pi,\pi]$). For example

$$f(x) = g(x) = \begin{cases} 0 & 0 \\ x^{-0.3} & x\neq 0, \end{cases}$$

then

$$\int_{-\pi}^{\pi} (fg)^2 = \int_{-\pi}^{\pi} x^{-1.2}dx \to \infty.$$

A consequence of more importance is, if $\{f(n)\},\{g(n)\}\in l^2$, its convolution needn't be in l^2. For example,

$$f(n) = g(n) = \begin{cases} \frac{1}{|n|^{0.6}} & n\neq 0 \\ 0 & n=0, \end{cases}$$

then

$$h(n)=\sum_{k} f(k)\,g(n-k)=\sum_{\substack{k\neq 0\\ k\neq n}} \frac{1}{|k(n-k)|^{.6}}\ .$$

For n>2, n-1>n/2 and when 0<k<n,

$$k(n-k)\le n^2,$$

then

$$h(n)\ge \sum_{0<k<n} \frac{1}{|k(n-k)|^{.6}} \ge \sum_{0<k<n} \frac{1}{n^{1.2}} = \frac{n-1}{n^{1.2}} \ge \frac{n}{2n^{1.2}} = \frac{1}{2n^{.2}}\ .$$

Clearly, {h(n)} is not in l^2.

LEMMA 5. If $\{f(n)\}\in l^2$, $\{g(n)\}\in l^1$, then the discrete convolution of f and g satisfies $\{h(n)\}\in l^2$ and $\hat{h}^*=\hat{f}^*\hat{g}^*$ in $L^2[-\pi,\pi]$.

Proof. Define

$$h_N(n)=\sum_{k=-N}^{N} f(k)\,g(n-k)\,,$$

then

$$\hat{h}_N^*(\omega)=\sum_{n} h_N(n)\,e^{-i\omega n}$$

$$=\sum_n \left(\sum_{k=-N}^{N} f(k)\, g(n-k)\right) e^{-i\omega n}$$

$$=\sum_n \sum_{k=-N}^{N} f(k)\, e^{-i\omega k} g(n-k)\, e^{-i\omega (n-k)} .$$

By repeated use of Lemma 0, the above expression is equal to

$$\sum_{k=-N}^{N} f(k)\, e^{-i\omega k} \sum_n g(n-k)\, e^{-i\omega (n-k)} = \sum_{k=-N}^{N} f(k)\, e^{-i\omega n} \cdot \hat{g}^*(\omega) .$$

Since

$$\sum_{k=-N}^{N} f(k)\, e^{-i\omega k} \to \hat{f}^*(\omega) \quad in\ L^2[-\pi,\pi] ,$$

and $\hat{g}^*(w)$ is bounded (lemma 3), using Lemma 1 gives

$$\lim_{N\to\infty} \hat{h}_N^*(\omega) \to \hat{f}^*(\omega)\, \hat{g}^*(\omega) \quad in\ L^2[-\pi,\pi] .$$

For any N>0, define

$$f^N(k) = \begin{cases} f(k) & |k|>N \\ 0 & |k|\le N \end{cases} ,$$

then

$$h(n) - h_N(n) = \sum_{|k|>N} f(k)\, g(n-k) = \sum_k f^N(k)\, g(n-k) .$$

By lemma 4,

$$\sum_n |h(n) - h_N(n)|^2 \le \left[\sum_n |g(n)|\right]^2 \sum_n |f^N(n)|^2$$

By assumption $\{f(n)\}\epsilon l^2$, so as $N\to\infty$,

$$\sum_n |f^N(n)|^2 \to 0.$$

Hence $\{h_N(n)\}$ converges to $\{h(n)\}$ in l^2, by lemma 2′,

$$\hat{h}_N^*(\omega)\to\hat{h}^*(\omega) \text{ in } L^2[-\pi,\pi].$$

Therefore,

$$\hat{h}^*(\omega)=\lim_{N\to\infty}\hat{h}_N^*(\omega)=\hat{f}^*(\omega)\,\hat{g}^*(\omega) \text{ in } L^2[-\pi,\pi].$$

By lemma 3, $\hat{g}^*$ is bounded. Thus $\hat{h}^*\epsilon L^2[-\pi,\pi]$ and therefore $\{h(n)\}\epsilon l^2$. ▮

DISCUSSION. Any function of L^p is actually an equivalence class of functions. If a sampling theorem holds for one representative of the class, it might not hold for another one. For example, if a non-zero function f satisfies

$$f(t)=\sum_n f(n)\,S(t-n)$$

then there is at least one integer k such that $f(k)\neq 0$. To

$$f_e(x)=\begin{cases} f(x) & x\neq k \\ 0 & x=k \end{cases}$$

use of the same sampling theorem will lead to the contradictory result

$$f_e(x)=\sum_n f_e(n)\,S(x-n)=\sum_n f(n)\,S(x-n)-f(k)\,S(x-k)$$

The reason the sampling theorem doesn't work in this case is because the wrong representative of the equivalence class was chosen.

The sampling theorem resulting from a wavelet subspace created by multiresolution analysis is described next. ϕ is a scaling function generating a Multiresolution Analysis $\{V_m\}$. In the space V_0, the most important property is that $\{\phi(t-k)\}$, $k\epsilon Z$ forms a basis of the space $V_0 \subset L^2(R)$. Then for any $f\epsilon V_0$, there exists a unique sequence $\{a_k\}$ in l^2 such that

$$f(t)=\sum_k a_k\phi(t-k) \quad in\ L^2(R).$$

In this section, the function f will always denote some member of V_0 and $\{a_k\}$ denotes the l^2 sequence associated with f by the above equation.

THEOREM 3.1. If ϕ is a real scaling function such that

(i) $\{\phi(n)\}\epsilon l^1$, $\hat{\phi}(w)$ is bounded and $\hat{\phi}(w)=O(w^{-1/2-\varepsilon})$ as $w\to\pm\infty$ for some $\varepsilon>0$;

(ii) $\hat{\phi}^*(\omega)=\Sigma_n\phi(n)e^{-i\omega n}\neq 0$, for all $\omega\epsilon R$;

(iii) $f\epsilon V_0$ and $f(n)=\Sigma_k a_k\phi(n-k)$

then

$$f(t)=\sum_{n=-\infty}^{+\infty} f(n)S(t-n)$$

in L^2, where S(t) is defined by

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega).$$

Proof. By assumption

$$f(n)=\sum_k a_k\phi(n-k).$$

Since $\{a_n\}\in l^2$ and $\{\phi(n)\}\in l^1$, using lemma 5 gives

$$\hat{f}^*(\omega)=\hat{a}^*(\omega)\hat{\phi}^*(\omega) \text{ in } L^2[-\pi,\pi].$$

Define

$$f_N(t)=\sum_{k=-N}^{N} a_k\phi(t-k),$$

then

$$\hat{f}_N(\omega)=\sum_{k=-N}^{N} a_k e^{-i\omega k}\cdot\hat{\phi}(\omega).$$

By assumption, $\hat{\phi}(w)$ is bounded and $\hat{\phi}(w)=O(w^{-1/2-\epsilon})$ as $w\to\pm\infty$, suppose that

$$|\hat{\phi}(\omega)|\le M \text{ and } \hat{\phi}(\omega)=O(\omega^{-1/2-\epsilon}) \text{ for } |\omega|\ge\pi W,$$

where W is some odd number. Use the fact that

$$\hat{a}^*(\omega)=\sum_{n=-N}^{N} a_n e^{-i\omega n}$$

is 2π-periodic, and

$$|\hat{\phi}(\omega)|^2 = O(\omega^{-1-2\epsilon}) = O(k^{-1-2\epsilon}) \quad for\ |\omega| \geq W\pi\ and\ (2k-1)\pi \leq \omega < (2k+1)\pi,$$

then

$$\int_{-\infty}^{\infty} |\hat{f}_N(\omega) - \hat{a}^*(\omega)\hat{\phi}(\omega)|^2 d\omega = \sum_{k=-\infty}^{\infty} \int_{(2k-1)\pi}^{(2k+1)\pi} |\hat{\phi}(\omega)|^2 |\hat{a}^*(\omega) - \sum_{n=-N}^{N} a_n e^{-i\omega n}|^2 d\omega$$

$$\leq [M^2 W + \sum_{k\neq 0} O(k^{-1-2\epsilon})] \int_{-\pi}^{\pi} |\hat{a}^*(\omega) - \sum_{n=-N}^{N} a_n e^{-i\omega n}|^2 d\omega$$

Since $\{a_k\} \in l^2$

$$\sum_{k=-N}^{N} a_k e^{i\omega k} \to \hat{a}^*(\omega) \quad in\ L^2[-\pi, \pi],$$

Therefore

$$\hat{f}_N(\omega) \to \hat{a}^*(\omega)\hat{\phi}(\omega) \quad in\ L^2(R).$$

Clearly $f_N \to f$ in $L^2(R)$, by lemma 2,

$$\hat{f}(\omega) = \lim_{N\to\infty} \hat{f}_N(\omega) = \hat{a}^*(\omega)\hat{\phi}(\omega) \quad in\ L^2(R).$$

By assumption $\hat{\phi}^*(\omega) \neq 0$, using Lemma 3 gives

$$\hat{\phi}^*(\omega) \quad \geq m$$

for some m>0. Thus $\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$ is well defined, $|\hat{S}(w)|\leq M/m$ and $\hat{S}(w)=O(w^{-1/2-\varepsilon})$ for $|w|\geq W\pi$, therefore

$$\int_{-\infty}^{\infty}|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)-\hat{f}(\omega)|^2 d\omega$$

$$=\int_{-\infty}^{\infty}|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)-\hat{a}^*(\omega)\hat{\phi}(\omega)|^2 d\omega$$

$$=\sum_{k}\int_{(2k-1)\pi}^{(2k+1)\pi}|\hat{S}(\omega)|^2|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}-\hat{a}^*(\omega)\hat{\phi}^*(\omega)|^2 d\omega$$

$$\leq[\frac{M^2W}{m^2}+\sum_{k\neq 0} O(k^{-1-2\varepsilon})]\int_{-\pi}^{\pi}|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}-\hat{a}^*(\omega)\hat{\phi}^*(\omega)|^2 d\omega$$

$$=M'\int_{-\pi}^{\pi}|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}-\hat{f}^*(\omega)|^2 d\omega,$$

where M′ is some positive number. Because

$$\int_{-\pi}^{\pi}|\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}-\hat{f}^*(\omega)|^2 d\omega$$

can be arbitrarily small as $N\to\infty$,

$$\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)\to\hat{f}(\omega) \ \ in\ L^2(R).$$

Since $\phi(t)$ is an element of V_0, it is in $L^2(R)$, hence $\hat{\phi}(\omega)\in L^2(R)$, the inequality

$$\|\hat{S}\|_{L^2(R)}\leq\frac{1}{m}\|\hat{\phi}\|_{L^2(R)}$$

makes $\hat{S}\epsilon L^2(R)$, so $S\epsilon L^2$ and if $g_N(t)$ is defined as

$$g_N(t) = \sum_{n=-N}^{N} f(n)S(t-n)$$

then $\hat{g}_N\epsilon L^2$. Thus

$$\hat{g}_n(\omega) = \sum_{n=-N}^{N} f(n)e^{-i\omega n}\hat{S}(\omega) \to \hat{f}(\omega)$$

in $L^2(R)$. By lemma 2,

$$f(t) = \lim_{N\to\infty} g_n(t) = \sum_n f(n)S(t-n)$$

in $L^2(R)$. ∎

Henceforth, if the function S exists, it will be called the **sampling kernel** related to scaling function ϕ. Many known scaling functions satisfy conditions (i) and (ii) in the above theorem. However, the "proper choice condition" $f(n)=\Sigma_k a_k\phi(n-k)$, which is needed for the sampling theorem to be valid in the sense of L^2 is usually very hard to check. On the other hand, if the choice of functions in V_0 is restricted, and ϕ and S are defined "properly", the following theorem and it's corollaries show that the "proper choice condition" holds and the sampling theorem is valid both in the L^2 sense and L^∞ sense (uniform convergence).

THEOREM 3.2. If ϕ is a real scaling function such that

(i) $\{\phi(n)\}\epsilon l^1$, $\hat{\phi}(w)\epsilon L^1$;

(ii) $\hat{\phi}^*(\omega)=\Sigma_n \phi(n)e^{-i\omega n}\neq 0$, for all $\omega\in R$;

(iii) $f\in V_0$, $\{a_n\}\in l^1$;

Define

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$$

and suppose $\phi(t)$, $S(t)$ and $f(t)$ are defined using the inverse Fourier transform formula

$$g(t)=\mathcal{F}^{-1}(g)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{g}(\omega)\,e^{i\omega t}d\omega$$

Then

$$f(t)=\sum_{n=-\infty}^{+\infty} f(n)\,S(t-n)$$

where the convergence is both in $L^2(R)$ and $L^\infty(R)$.

Proof. Define

$$f_N(t)=\sum_{k=-N}^{N} a_k\phi(t-k)\ .$$

Taking Fourier transform of both side gives

$$\hat{f}_N(\omega)=\sum_{k=-N}^{N} a_k e^{-i\omega k}\cdot\hat{\phi}(\omega)\,,$$

hence

$$\int_{-\infty}^{\infty}|\hat{f}_N(\omega)-\hat{a}^*(\omega)\hat{\phi}(\omega)|^2 d\omega=\int_{-\infty}^{\infty}|\hat{\phi}(\omega)|^2|\sum_{|n|>N} a_n e^{-i\omega n}|^2 d\omega.$$

Because $\{a_k\}\in l^1$, and

$$|\sum_{|n|>N} a_n e^{-i\omega n}|\le\sum_{|n|>N}|a_k|,$$

$$|\sum_{|n|>N} a_n e^{-i\omega n}|$$

can be arbitrarily small as $N\to\infty$. Therefore

$$\hat{f}_N(\omega)\to\hat{a}^*(\omega)\hat{\phi}(\omega)\ in\ L^2(R).$$

Clearly $f_N\to f$ in $L^2(R)$, by lemma 2,

$$\hat{f}(\omega)=\lim_{N\to\infty}\hat{f}_N(\omega)=\hat{a}^*(\omega)\hat{\phi}(\omega)\ in\ L^2(R).$$

So

$$\hat{f}(\omega)=\hat{a}^*(\omega)\hat{\phi}(\omega)\ a.e.$$

Since $\{a_k\}\in l^1$, $\hat{a}^*(w)$ is bounded (lemma 3). Because $\hat{\phi}(w)\in L^1$, $\hat{f}(w)\in L^1$, so f defined as

$$f(t)=\mathcal{F}^{-1}(f)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{f}(\omega)e^{i\omega t}d\omega$$

is uniformly continuous (Theorem B1). By the same reason ϕ is uniformly continuous. So for any $\varepsilon>0$, there exists $\eta>0$ such that for $|\delta|<\eta$, $|\phi(t+\delta)-\phi(t)|<\varepsilon$ for any $t\in R$. therefore

$$\left|\sum_n a_n\phi(t+\delta-n)-\sum_n a_n\phi(t-n)\right|\le\sum_n\left|a_n[\phi(t+\delta-n)-\phi(t-n)]\right|\le\sum_n|a_n|\cdot\varepsilon .$$

So $\Sigma_n a_n\phi(t-n)$ is also continuous. Hence

$$f(t)=\sum_k a_k\phi(t-k) \quad in\ L^2$$

and both sides are continuous, thus they have to be equal everywhere. Otherwise suppose they are not equal at t_0, then they will not be equal within a small interval containing t_0, this makes them not equal in the L^2 sense.

Since the above equation holds everywhere, it follows that

$$f(n)=\sum_k a_k\phi(n-k) .$$

Since $\{a_k\}\in l^1$, $\{\phi(k)\}\in l^1$, then

$$\sum_n|f(n)|\le\sum_n\sum_k|a_k\phi(n-k)|=\sum_k|a_k|\sum_n|\phi(n-k)|=\sum_k|a_k|\sum_n|\phi(n)|.$$

So $\{f(n)\}\in l^1$ and

$$\sum_n\sum_k a_k\phi(n-k)\,e^{-i\omega n}$$

converge absolutely and the order of summation can be interchanged

$$\sum_n \sum_k a_k \phi(n-k) e^{-i\omega n} = \sum_k a_k e^{-i\omega k} \sum_n \phi(n-k) e^{-i\omega(n-k)}$$

$$= \sum_k a_k e^{-i\omega k} \sum_n \phi(n) e^{-i\omega n}$$

Thus

$$\hat{f}^*(\omega) = \hat{a}^*(\omega) \hat{\phi}^*(\omega) \text{ pointwise}$$

By definition

$$\hat{S}(\omega) = \hat{\phi}(\omega) / \hat{\phi}^*(\omega),$$

therefore

$$\int_{-\infty}^{\infty} \left| \sum_{n=-N}^{N} f(n) e^{-i\omega n} \hat{S}(\omega) - \hat{f}(\omega) \right|^2 d\omega$$

$$= \int_{-\infty}^{\infty} \left| \sum_{n=-N}^{N} f(n) e^{-i\omega n} \hat{S}(\omega) - \hat{a}^*(\omega) \hat{\phi}(\omega) \right|^2 d\omega$$

$$= \int_{-\infty}^{\infty} |\hat{S}(\omega)|^2 \left| \sum_{n=-N}^{N} f(n) e^{-i\omega n} - \hat{a}^*(\omega) \hat{\phi}^*(\omega) \right|^2 d\omega$$

$$= \int_{-\infty}^{\infty} |\hat{S}(\omega)|^2 \left| \sum_{n=-N}^{N} f(n) e^{-i\omega n} - \hat{f}^*(\omega) \right|^2 d\omega$$

$$=\int_{-\infty}^{\infty}|\hat{S}(\omega)|^2\Big|\sum_{|n|>N} f(n)\, e^{-i\omega n}\Big|^2 d\omega$$

By the fact $\{f(n)\}\in l^1$,

$$\Big|\sum_{|n|>N} f(n)\, e^{-i\omega n}\Big|\to 0 \ \ as\ N\to\infty.$$

Hence

$$\sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)\to\hat{f}(\omega)\ \ in\ L^2(R)$$

By assumption $\hat{\phi}^*(\omega)\neq 0$ and $\{\phi(n)\}\in l^1$, using lemma 3 gives

$$\hat{\Phi}^*(\omega)\ \ \geq m$$

for some m>0. Thus $|\hat{S}(w)|\leq|\hat{\phi}(\omega)|/m$.

Since $\phi(t)$ is an element of V_0, it is in $L^2(R)$, hence $\hat{\phi}(\omega)\in L^2(R)$, and

$$\|\hat{S}\|_{L^2(R)}\leq\frac{1}{m}\|\hat{\Phi}\|_{L^2(R)}$$

makes $\hat{S}\in L^2(R)$, so $S\in L^2(R)$ and if $g_N(t)$ is defined as

$$g_N(t)=\sum_{n=-N}^{N} f(n)\, S(t-n)$$

then $\hat{g}_N \in L^2(R)$. Clearly

$$\hat{g}_n(\omega) = \sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega) \to \hat{f}(\omega)$$

in $L^2(R)$. By lemma 2,

$$f(t) = \lim_{N\to\infty} g_n(t) = \sum_{n} f(n)\, S(t-n)$$

in $L^2(R)$.

Since

$$\hat{f}(\omega) = \hat{a}^*(\omega)\hat{\phi}(\omega) \quad a.e.$$

and

$$\hat{a}^*(\omega)\hat{\phi}(\omega) = \hat{a}^*(\omega)\hat{\phi}^*(\omega)\hat{S}(\omega) = \hat{f}^*(\omega)\hat{S}(\omega)$$

everywhere, hence

$$\int_{-\infty}^{\infty} \hat{f}(\omega) - \hat{f}^*(\omega)\hat{S}(\omega) \,|\, d\omega = \int_{-\infty}^{\infty} \hat{f}(\omega) - \hat{a}^*(\omega)\hat{\phi}(\omega) \,|\, d\omega = 0.$$

Also $|\hat{S}(w)| \le |\hat{\phi}(\omega)|/m$, then $\hat{S}(w) \in L^1(R)$. Therefore

$$\left| f(t) - \sum_{n=-N}^{N} f(n)\, S(t-n) \right|$$

$$= \frac{1}{2\pi} \left| \int_{-\infty}^{\infty} [\hat{f}(\omega) - \sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)]\, e^{i\omega t} d\omega \right.$$

$$\le \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}(\omega) - \sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\hat{S}(\omega)\,|\, d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\hat{f}^{*}(\omega) - \sum_{n=-N}^{N} f(n)\, e^{-i\omega n}\,|\,|\hat{S}(\omega)|\, d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} |\sum_{|n|>N} f(n)\, e^{-i\omega n}\,|\,|\hat{S}(\omega)\,|\, d\omega$$

$$\le \frac{1}{2\pi}\sum_{|n|>N} |f(n)|\int_{-\infty}^{\infty} |\hat{S}(\omega)|\, d\omega \to 0 \text{ as } N\to\infty.$$

This proves the uniform convergence. ∎

COROLLARY 1. If ϕ is a real scaling function such that

(i) $\{\phi(n)\}\in l^1$, $\hat{\phi}(w)\in L^1$;

(ii) $\hat{\phi}^*(\omega)=\sum_n \phi(n)e^{-i\omega n}\neq 0$, for all $\omega\in R$;

(iii) $f\in V_0$, $\{a_n\}\in l^1$;

Define

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$$

and suppose $\phi(t)$, $S(t)$ and $f(t)$ are all continuous and in $L^1(R)$. Then

$$f(t)=\sum_{n=-\infty}^{+\infty} f(n)\,S(t-n)$$

where the convergence is both in $L^2(R)$ and $L^\infty(R)$.

Proof. The corollary is proved by simply noting the fact that if g(t) continuous, g and $\hat{g}$ are in $L^1(R)$, then by theorem B2 (see appendix B), g(t) can be recovered pointwise by its inverse Fourier transform

$$g(t) = \mathcal{F}^{-1}(g) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{g}(\omega)\, e^{i\omega t}d\omega$$

When ϕ is an o.n. scaling function, then

$$a_k = \int_{-\infty}^{\infty} f(x)\,\phi(x-k)\,dx.$$

Therefore

$$\sum_{k=-N}^{N}|a_k| = \sum_{k=-N}^{N}\left|\int_{-\infty}^{\infty} f(x)\,\phi(x-k)\,dx\right|$$

$$\leq \sum_{k=-N}^{N}\int_{-\infty}^{\infty}|f(x)\,\phi(x-k)|\,dx$$

$$\leq \int_{-\infty}^{\infty}\sum_{k=-N}^{N}|f(x)\,\phi(x-k)|\,dx$$

$$\leq \int_{-\infty}^{\infty}|f(x)|\cdot\sum_{k}|\phi(x-k)|\,dx,$$

thus

$$\sum_k |a_k| \le \int_{-\infty}^{\infty} |f(x)| \cdot \sum_k |\phi(x-k)| \, dx.$$

Similarly, notice

$$a_{-k} = \int_{-\infty}^{\infty} f(x)\phi(x+k)\,dx = \int_{-\infty}^{\infty} f(x-k)\phi(x)\,dx,$$

$$\sum_k |a_k| \le \int_{-\infty}^{\infty} |\phi(x)| \cdot \sum_k |f(x-k)| \, dx.$$

Therefore if one of f(x) or $\phi(x) \in L^1$, and the other is $O(x^{-1-\varepsilon})$ as $x \to \pm\infty$, then $\{a_k\} \in l^1$. Clearly, just use the above inequalities, different conditions making $\{a_k\} \in l^1$ are also possible. Additionally, by using a similar analysis, conditions that force $\{a_k\} \in l^1$ can be obtained with semi-orthonormal and non-orthogonal scaling functions.

The above discussion leads to the following corollary.

COROLLARY 2. If ϕ is a continuous o.n. scaling function such that

(i) $\{\phi(n)\} \in l^1$, $\phi(t) = O(t^{-1-\varepsilon})$ as $t \to \pm\infty$ and $\hat{\phi}(w) \in L^1$;

(ii) $\hat{\phi}^*(\omega) = \sum_n \phi(n) e^{-i\omega n} \ne 0$, for all $\omega \in R$;

Define

$$\hat{S}(\omega) = \hat{\phi}(\omega) / \hat{\phi}^*(\omega)$$

and suppose S(t) and f(t) are all continuous and all in $L^1(R)$. Then for $f \in V_0$,

$$f(t) = \sum_{n=-\infty}^{+\infty} f(n)\, S(t-n)$$

where the convergence is both in $L^2(R)$ and $L^\infty(R)$.

3.2. Continuous Multiresolution Analysis and Sampling Theorems in U_α

As the example 2.2. shows

$$\phi(t)=\frac{\sin\pi t}{\pi t}$$

is the scaling function with V_0 being functions bandlimited to $[-\pi,\pi]$. Correspondingly, V_n is the space with functions bandlimited to $[-2^n\pi,2^n\pi]$. In order to cover the "scale gap" in the discrete MRA, new spaces are to be "filled in".

First an identification is made between V_n with U_{2^n}. Use the fact that $x\mapsto 2^x$ is a one-to-one and onto mapping between R and $(0,+\infty)$. A continuous scaling scheme, henceforth called continuous Multiresolution analysis can be found by working with U_α with α nonnegative.

Definition 3.5. A collection of subspaces U_α, $\alpha\epsilon R^+$ (R^+ denote the set of all positive real numbers) of $L^2(R)$ is said to form a **continuous multiresolution analysis** if

(i) $\{U_{2^n}\}$ forms a discrete MRA

(ii) $f(x)\epsilon U_1 \Leftrightarrow f(\alpha x)\epsilon U_\alpha$, $\alpha\epsilon R$;

The function ϕ which generates discrete MRA $\{U_{2^n}\}$ will be called the scaling function which generates the continuous multiresolution analysis U_α, $\alpha \in R^+$. Clearly, every discrete MRA uniquely determines a continuous MRA.

Because $\{\phi(\alpha t-k)\}$ is a Riesz basis of U_α (see proposition 3.2. in section 3.3), for any $f \in V_\alpha$, there exists $\{a_k\}$ such that

$$f(t) = \sum_k a_k \phi(\alpha t-k) \text{ in } L^2(R).$$

Hence in the space U_α, the "proper choice condition" becomes

$$f(n/\alpha) = \sum_k a_k \phi(n-k).$$

The following is the extension of theorem 3.1. to space U_α.

THEOREM 3.3. If ϕ is real scaling function such that

(i) $\{\phi(n)\} \in l^1$, $\hat{\phi}(w)$ is bounded and $\hat{\phi}(w)=O(w^{-1/2-\varepsilon})$ as $w \to \pm\infty$ for some $\varepsilon>0$;

(ii) $\hat{\phi}^*(\omega)=\Sigma_n \phi(n)e^{-i\omega n} \neq 0$, for all $\omega \in R$;

(iii) $f \in U_\alpha$ and $f(n/\alpha)=\Sigma_k a_k \phi(n-k)$

then

$$f(t) = \sum_{n=-\infty}^{+\infty} f(n/\alpha) S(\alpha t-n)$$

in L^2, where $S(t)$ is defined as the inverse Fourier transform of

$$\hat{S}(\omega) = \hat{\phi}(\omega)/\hat{\phi}^*(\omega).$$

Proof. By the properties of the Multiresolution analysis, $f(\cdot)\in U_\alpha$ means $f(\cdot/\alpha)\in U_1=V_0$. Define

$$f_\alpha(t)=f(t/\alpha)$$

then $f_\alpha\in V_0$ and $f_\alpha(n)=\Sigma_k a_k\phi(n-k)$. Using theorem 3.1. gives

$$f(t/\alpha)=f_\alpha(t)=\sum_n f_\alpha(n)S(t-n) \text{ in } L^2.$$

That is equivalent to

$$f(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)S(\alpha t-n) \text{ in } L^2. \qquad ▮$$

Extension of theorem 3.2. is formulated as follows.

THEOREM 3.4. If ϕ is a real scaling function such that

(i) $\{\phi(n)\}\in l^1$, $\hat{\phi}(w)\in L^1$;

(ii) $\hat{\phi}^*(\omega)=\Sigma_n\phi(n)e^{-i\omega n}\neq 0$, for all $\omega\in R$;

(iii) $f\in U_\alpha$, $\{a_n\}\in l^1$;

Define

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$$

and suppose $\phi(t)$, $S(t)$ and $f(t)$ are defined using the inverse Fourier transform formula

$$g(t)=\mathcal{F}^{-1}(g)=\frac{1}{2\pi}\int_{-\infty}^{\infty}\hat{g}(\omega)\,e^{i\omega t}d\omega$$

Then

$$f(t)=\sum_{n} f(n/\alpha)\, S(\alpha t-n)$$

where the convergence is both in $L^2(R)$ and $L^\infty(R)$.

Extension of the corollaries of theorem 3.2. can be formulated similarly.

3.3. Aliasing Error Analysis

Consider the Whittaker-Shannon sampling theorem. Only when the signal is strictly bandlimited perfect reconstruction is possible. However such signals are not realizable in real life since the only signal both timelimited and bandlimited is zero, and all real world signals are timelimited. Thus in real world applications there is aliasing. Additional error may occur because of imperfect timing and jitter. Generally speaking, to any sampling theorem, there are four types of errors, namely, truncation error, jitter error, round-off error and aliasing error. Numerous paper have been written describing these errors.

The sampling theory developed above shows that if the wavelet is chosen properly, then for $f \in U_\alpha$,

$$f(t)=\sum_n f(n/\alpha)\,S(\alpha t-n)$$

When $f \notin U_\alpha$, the above sampling theorem will not hold exactly. Define the sampling series of f to be

$$g(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\,S(\alpha t-n)$$

The purpose of aliasing error analysis is to determine how close f and g will be when $f \notin U_\alpha$. First, consider the case when $f \notin V_0 = U_1$, here $\alpha=1$.

THEOREM 3.5. Suppose V_0 is a wavelet subspace with an associated sampling theorem holding and its sampling kernel S satisfies the assumption that the partial sums of

$$\psi(t,\omega) = \sum_{n=-\infty}^{+\infty} e^{i\omega n} S(t-n)$$

are uniformly bounded in t and ω. Suppose $f \in L^2$ is written as $f = f_0 + h$, where $f_0 \in V_0$ and

$$f(t) = \sum_{n} f(n) S(t-n) \text{ pointwise,}$$

while h can be written as

$$h(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} d\beta(\omega)$$

where $\beta(\omega)$ is of bounded variation from $-\infty$ to $+\infty$, then the sampling series g of f converges pointwise and

$$\|f-g\|_\infty \leq C \cdot \bigvee_{-\infty \leq \omega \leq +\infty} \beta(\omega)$$

where the constant C depends on the type of wavelet employed.

Proof.

$$g(t) = \sum_{n=-\infty}^{+\infty} f(n) S(t-n)$$

$$=\sum_{n=-\infty}^{+\infty}(f_0(n)+h(n))S(t-n)$$

$$=\sum_{n=-\infty}^{+\infty}f_0(n)S(t-n)+\sum_{n=-\infty}^{+\infty}h(n)S(t-n)$$

$$=f_0(t)+\sum_{n=-\infty}^{+\infty}h(n)S(t-n)$$

Denote

$$h_s(t)=\sum_n h(n)S(t-n)$$

Notice

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\omega t}d\beta(\omega)$$

and in particular

$$h(n)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{i\omega n}d\beta(\omega)$$

So

$$h_s(t)=\frac{1}{2\pi}\sum_{n=-\infty}^{+\infty}\int_{-\infty}^{+\infty}e^{i\omega n}d\beta(\omega)\cdot S(t-n)$$

By assumption, the partial sums of $\psi(t,\omega)$

$$\psi_K(t,\omega)=\sum_{n=-K}^{K}e^{i\omega n}S(t-n)$$

are uniformly bounded, say with bound M. Also by assumption $\beta(\omega)$ is of bounded variation, then

$$\int_{-\infty}^{+\infty} |\psi(t,\omega)\, d\alpha(\omega)| \leq M \cdot \bigvee_{-\infty\leq\omega\leq+\infty} \beta(\omega)$$

Apply Lebesgue's Dominated Convergence Theorem, the order of summation and integration can be interchanged in the calculation of $h_s(t)$. Thus

$$h_s(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} e^{i\omega n} d\beta(\omega) \cdot S(t-n) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \psi(t,\omega)\, d\beta(\omega)$$

and

$$f(t)-g(t) = h(t)-h_s(t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} (e^{i\omega t}-\psi(t,\omega))\, d\beta(\omega)$$

$$\|f-g\|_\infty \leq C \cdot \bigvee_{-\infty\leq\omega\leq-\infty} \beta(\omega)$$

where C can be chosen as

$$C = \frac{1}{2\pi}\|\psi(t,\omega)-e^{i\omega t}\|_\infty \qquad \blacksquare$$

Using the scaling argument, a similar result for the space U_α is formulated as follows.

COROLLARY. Suppose the wavelets subspaces $\{U_\alpha\}$ with an associated sampling theorem holding and sampling kernel S with the partial sum of

$$\psi(t,\omega)=\sum_{n=-\infty}^{+\infty} e^{i\omega n}S(t-n)$$

uniformly bounded in t and ω. Suppose $f\epsilon L^2$ is written as $f=f_\alpha+h$, where $f_\alpha \epsilon U_\alpha$ and

$$f(t)=\sum_n f(n/\alpha)\, S(\alpha t-n) \text{ pointwise,}$$

while h can be written as

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t} d\beta(\omega)$$

where $\beta(w)$ is of bounded variation from $-\infty$ to $+\infty$, then the sampling series g of f converges and

$$\|f-g\|_\infty \leq C' \cdot \bigvee_{-\infty\leq\omega\leq+\infty} \beta(\omega)$$

where the constant C′ depends on the individual wavelets.

Proof. Use the scale change method used in the proof of theorem 3.3. of this chapter. Define

$$g^*(t)=g(t/\alpha),\quad f^*(t)=f(t/\alpha),\quad f_0^*(t)=f_\alpha(t/\alpha),\quad h^*(t)=h(t/\alpha)$$

then $f_0^* \epsilon V_0$ so

$$g^*(t)=\sum_{n=-\infty}^{+\infty} f^*(n)\, S(t-n)$$

$$= f_0^*(t) + \sum_{n=-\infty}^{+\infty} h^*(n) S(t-n)$$

and by

$$h^*(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{i\omega t/\alpha} d\beta(\omega)$$

Therefore

$$h_s^*(t) = \sum_n h^*(n) S(t-n) = \frac{1}{2\pi} \sum_{n=-\infty}^{+\infty} \int_{-\infty}^{+\infty} e^{i\omega n/\alpha} d\beta(\omega) \cdot S(t-n)$$

Notice

$$\psi_K(t, \omega/\alpha) = \sum_{n=-K}^{K} e^{i\omega n/\alpha} S(t-n)$$

is by assumption uniformly bounded, say with bound M. Also by assumption $\beta(\omega)$ is of bounded variation, then

$$\int_{-\infty}^{+\infty} |\psi(t, \omega) d\beta(\omega)| \leq M \cdot \bigvee_{-\infty \leq \omega \leq +\infty} \beta(\omega)$$

Apply Lebesgue's Dominated Convergence Theorem, the order of summation and integration can be interchanged in the calculation of $h_s^*(t)$. Thus

$$h_s^*(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} e^{i\omega n/\alpha} d\beta(\omega) \cdot S(t-n) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \psi(t, \omega/\alpha) d\beta(\omega)$$

$$f^*(t)-g^*(t)=h^*(t)-h_s^*(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}[e^{i\omega t/\alpha}-\psi(t,\omega/\alpha)]\,d\beta(\omega)$$

This is the same as

$$f(t)-g(t)=f^*(\alpha t)-g^*(\alpha t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}(e^{i\omega t}-\psi(\alpha t,\omega/\alpha))\,d\beta(\omega)$$

So

$$\|f-g\|_\infty\le C'\cdot\bigvee_{-\infty\le\omega\le-\infty}\beta(\omega)$$

where C′ can be chosen as

$$C'=\frac{1}{2\pi}\|\psi(\alpha t,\omega/\alpha)-e^{i\omega t}\|_\infty \qquad \blacksquare$$

Notice that the constant C′ might be different for different U_α. Using the fact that $\psi(t,\omega)$ is uniformly bounded and by simply choosing C to be

$$C=\frac{1}{2\pi}(\|\psi(t,\omega)\|_\infty+1)$$

a uniform bound for all U_α is obtained.

In using the above theorems, the condition on $\psi(t,\omega)$ seems to be a little hard to determine. An important characterization of $\psi(t,\omega)$ will be presented by theorem 4.4.

In order to analyze $\beta(w)$ as $\alpha\to+\infty$, a detailed discussion about U_α is given below.

Proposition 3.1. If there is a discrete MRA with scaling function ϕ, then there is a corresponding continuous MRA with the same scaling function satisfying the following properties.

(a) $\mathrm{Clos}_{L^2}(\bigcup_{\alpha\in R} U_\alpha) = L^2(R)$;

(b) $\bigcap U_{\alpha\in R} = \{0\}$;

(c) $f(x)\in U_\alpha \Leftrightarrow f(\beta x)\in U_{\alpha\beta}$, $\alpha,\beta\in R$;

(d) $f(x)\in U_\alpha \Leftrightarrow f(x+1/\alpha)\in U_\alpha$, $\alpha\in R$.

Proof.

(a) Notice $U_\alpha \subset L^2(R)$ and $\bigcup_{\alpha\in R} U_\alpha \supset \bigcup_{\alpha=2^n, n\in Z} U_\alpha$, this makes (a) holds.

(b) By the fact $U_\alpha \supset \{0\}$ and $\bigcap U_{\alpha\in R} \subset \bigcap_{\alpha=2^n, n\in Z} U_\alpha$, so (b) holds.

(c) By (ii) of the continuous MRA axiom, for all $\alpha\in R$, $f(x)\in U_\alpha$ $\Leftrightarrow f(x/\alpha)\in U_1 \Leftrightarrow f(\beta x)=f(\beta\alpha\cdot x/\alpha)\in U_{\alpha\beta}$

(d) By (ii) in the continuous MRA axiom and (5°) in the discrete MRA axiom, for all $\alpha\in R$. $f(x)\in U_\alpha \Leftrightarrow f(x/\alpha)\in U_1 \Leftrightarrow$ $f(x/\alpha+1/\alpha)\in U_1 \Leftrightarrow f(x+1/\alpha)\in U_\alpha$. ▌

The relation of U_α and U_β for $\alpha,\beta\in R$ and $\alpha<\beta$ depend on the individual scaling function being chosen.

Example 3.1. Choose the scaling function to be $\chi_{[0,1)}$, then U_1 is the collection of all L^2 square pulse type functions with step width being 1, while $U_{3/4}$ the collection of all L^2 square pulse type function with step width being 3/4. In this case $U_{3/4} \not\subset U_1$.

Example 3.2. Choose the scaling function to be $\sin\pi t/\pi t$, then U_1 is Paley-Wiener space PW_π, that is the collection of all L^2 function with their Fourier transform bandlimited to $[-\pi,\pi]$. And U_α is the collection of all L^2 function bandlimited to $[-\alpha\pi,\alpha\pi]$. Here, $U_\alpha \subset U_\beta$ for all $\alpha,\beta \in R$ and $\alpha<\beta$.

Therefore in general, it is not true that $U_\alpha \subset U_\beta$ for $\alpha,\beta \in R$ and $\alpha<\beta$. For fixed α, the collection of subspaces $\{U_{\alpha 2^n}\}$, $n \in Z$ will have very similar properties as a discrete MRA, as the following theorem shows.

Proposition 3.2. If the function $\phi(x)$ is the scaling function of a discrete MRA $\{V_n\}$, then $\{V_n'\}$, where $V_n'=U_{\alpha 2^n}$ satisfies:

(1°) V_j' are closed and nested i.e. $\cdots \subset V_{-1}' \subset V_0' \subset V_1' \subset \cdots$;

(2°) $Clos_{L^2}(\bigcup_{j \in Z} V_j') = L^2(R)$;

(3°) $\bigcap V_{j \in Z}' = \{0\}$;

(4°) $f(x) \in V_j' \Leftrightarrow f(2x) \in V_{j+1}'$, $j \in Z$;

(5°) If $\{\phi(x-k)\}$ is a Riesz basis of V_0 with Riesz bounds A and B, $\{1/\sqrt{\alpha}\phi(\alpha x-k)\}$ is a Riesz basis of V_0' with Riesz bounds A and B.

Proof.

(1°) If $f_m(x) \in V_n'$ and $f_m(x) \to f(x)$ in $L^2(R)$, then by (c) of proposition 3.1., $f_m(x/\alpha) \in V_n$. The fact V_n is closed makes $f(x/\alpha) \in V_n$, using (c) of proposition 2.2 gives $f(x) \in V_n'$, so V_n' is closed. Also if $f(x) \in V_n'$, then $f(x/\alpha) \in V_n \subset V_{n+1}$, so $f(x) \in V_{n+1}'$, therefore $V_n' \subset V_{n+1}'$.

(2°) If $g(x) \in L^2$, define $h(x)=g(x/\alpha)$, then $h(x) \in L^2$. Since $\mathrm{Clos}_{L^2}(\bigcup_{j\in Z} V_j)= L^2$, for all $\varepsilon>0$, there exist a function $f(x) \in V_m$ for some $m \in Z$, such that $\|h-f\|_{L^2}=\int |h(x)-f(x)|^2 dx<\varepsilon$, so $\int |g(x)-f(\alpha x)|^2 d(\alpha x)=\int |h(\alpha x)-f(\alpha x)|^2 d(\alpha x)<\varepsilon$. Since $f(\alpha x) \in V_n'$, this means $\mathrm{Clos}_{L^2}(\bigcup_{j\in Z} V_j')= L^2(R)$.

(3°) If $f(x)\neq 0$ belongs to all V_n', then $f(x/\alpha)$ belongs to all V_n. Which is impossible. Thus $\bigcap_{n\in Z} V_n'=\{0\}$.

(4°) Follows directly from (c) of above proposition 3.1.

(5°) To $f(x) \in V_0'$, $f(x/\alpha) \in V_0$. Since $\{\phi(x-k)\}$ is a Riesz basis of V_0, there exists $\{c_k\} \in l^2$, such that $f(x/\alpha)=\sum c_k \phi(x-k)$. So $f(x)=\sum d_k \cdot 1/\sqrt{\alpha}\phi(\alpha x-k)$, where $d_k=\sqrt{\alpha} c_k$. So $\{d_k\} \in l^2$. Again use the fact that $\{\phi(x-k)\}$ is a Riesz basis of V_0, then there exists $B>A>0$, such that for any $\{c_k\} \in l^2$,

$$A\|\{c_k\}\|_{l^2}^2 \leq \|\sum_{k=-\infty}^{\infty} c_k \phi(x-k)\|_{L^2}^2 \leq B\|\{c_k\}\|_{l^2}^2$$

By a change of scale, this is equivalent to

$$A\|\{c_k\}\|_{l^2}^2 \leq \|\sum_{k=-\infty}^{\infty} c_k \frac{1}{\sqrt{\alpha}} \phi(\alpha x-k)\|_{L^2}^2 \leq B\|\{c_k\}\|_{l^2}^2$$

Hence $\{1/\sqrt{\alpha}\phi(\alpha x-k)\}$, $k \in Z$ is a Riesz basis for V_0' with Riesz bounds A and B. ▌

The only difference between $\{U_{\alpha 2^n}\}$ and $\{V_n\}$ is an scaling difference in (5°). By (ii) of the proposition 3.2., for any fixed α

$$\overline{\bigcup_{m\in Z} U_{\alpha 2^m}} = L^2$$

To any $f\in L^2$, it follows that

$$\lim_{m\to\infty}\|f-g\|_2=0$$

Since $\{U_{\alpha 2^n}\}$ is nested (see (i) of proposition 3.2.), this means that as $\alpha\to+\infty$,

$$\bigvee_{-\infty\leq\omega\leq\infty}\beta(\omega)\to 0.$$

3.4. Examples.

Several examples will be given where the sampling theorems and the error analysis theorems can be applied, and an example where the sampling theorem does not exist in the space of L^2 will also be noted.

EXAMPLE 3.3. Haar Wavelet.

The scaling function for the Haar wavelets is

$$\phi(t)=\chi_{[0,1)}$$

Thus

$$\hat{\phi}^{*}(\omega)=\sum_{n=-\infty}^{\infty}\phi(n)e^{-i\omega t}=\phi(0)=1$$

So

$$S(t)=\phi(t)=\chi_{[0,1)}$$

hence for $f\in U_{\alpha}$, the space of step functions with step α, the sampling theorem is

$$f(t)=\sum_{n=-\infty}^{\infty}f(n/\alpha)S(\alpha t-n)=f(n/\alpha)\quad when\ n/\alpha\le t<(n+1)/\alpha$$

Also

$$\psi(t,\omega)=\sum_{n=-\infty}^{+\infty} e^{i\omega t}S(t-n)=e^{i\omega[t]}$$

where [t] denote the largest integer no bigger than t. Since $e^{i\omega[t]}$ is of bounded variation on $[-\pi,\pi]$, by theorem B9, the partial sum of $\psi(t,\omega)$ are uniformly bounded. So if the other conditions specified in theorem 3.5. are satisfied, then

$$\|f-g\|_\infty \le C \bigvee_{-\infty\le\omega\le\infty} \beta(\omega)$$

where C can be chosen as

$$C=\frac{1}{\pi}\ge\frac{1}{2\pi}\|e^{i\omega/\alpha\cdot[\alpha t]}-e^{i\omega t}\|_\infty$$

Actually, by direct calculation

$$f(t)-g(t)=f(t)-f(n/\alpha) \quad \text{when } n/\alpha\le t<(n+1)/\alpha$$

So, theorem 3.5. does not seem to be of much help here. In fact, if S(t) is of compact support, the sampling series will be a finite sum. In that case, direct calculation of the sampling sum seems to be a better choice in deriving the aliasing errors. If S(t) doesn't have compact support, such as when $\hat{S}(\omega)$ has compact support, theorem 3.4. will be very useful as will be seen in the case of cardinal series and bandlimited sampling scaling functions of next chapter.

EXAMPLE 3.4. Cardinal series.

The scaling function is the Whittaker-Shannon sampling kernel

$$\phi(t)=\frac{\sin\pi t}{\pi t}$$

and U_α is a Paley-Wiener space with functions bandlimited in $[-\alpha\pi,\alpha\pi]$.

It is well known that

$$\hat{\phi}(\omega)=\begin{cases}1 & -\pi<\omega\leq\pi\\ 0 & otherwise\end{cases}$$

Hence

$$\sum_n \hat{\phi}(\omega+2\pi n)=1,\ \omega\in R$$

by theorem B11,

$$\hat{\phi}^*(\omega)=\sum_n \phi(n)\,e^{-i\omega n}=\sum_n \hat{\phi}(\omega+2\pi n)=1 \qquad \omega\in R$$

This makes $\hat{S}=\hat{\phi}$, $S=\phi$ and for $f\in U_\alpha$

$$f(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\,\frac{\sin\pi(\alpha t-n)}{\pi(\alpha t-n)}$$

which is exactly the Whittaker-Shannon sampling theorem.

As for dealing with the aliasing error, the following result holds.

THEOREM 3.6. If $f \in L^2$ and can be written in the form

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{i\omega t} d\beta(\omega)$$

where $\beta(\omega)$ is of bounded variation on $|\omega| \geq \sigma^-$, $\sigma=\alpha\pi$ and its sampling series is

$$g(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\, S(\alpha t-n)$$

then

$$\|f-g\|_\infty \leq \frac{1}{2\pi} \bigvee_{|\omega|\geq\alpha\pi^-} \beta(\omega)$$

Proof. If f_α is chosen to be the projection of f onto U_α, then

$$f_\alpha(t)=\frac{1}{2\pi}\int_{-\alpha\pi^+}^{\alpha\pi} e^{i\omega t} d\beta(\omega)$$

$$h(t)=\frac{1}{2\pi}\int_{|\omega|\geq\alpha\pi^-} e^{i\omega t} d\beta(\omega)$$

and

$$\psi(t,\omega)=\sum_{n=-\infty}^{+\infty} e^{i\omega n}\frac{\sin\pi(t-n)}{\pi(t-n)}$$

Using the fact

$$\psi(t,\omega)=\begin{cases} e^{i\omega t} & |\omega|<\pi \\ \cos\pi t & |\omega|=\pi \\ \psi(t,2\pi n+\omega) & |\omega|>\pi \end{cases}$$

and

$$f_\alpha(t) = \sum_n f(n/\alpha) S(\alpha t - n) \text{ pointwise,}$$

(see Giardina [1983]). Then simply by observing that $e^{i\omega t}$ is of bounded variation on $[-\pi,\pi]$, using theorem B9 gives that the partial sum of $\psi(t,\omega)$ is uniformly bounded. So by theorem 3.5,

$$\|f-g\|_\infty \le C \cdot \bigvee_{|\omega| \ge \alpha\pi^-} \beta(\omega)$$

Notice the fact norm of $\psi(t,\omega)$ and $e^{i\omega t}$ will never exceeds 1,

$$C = \frac{1}{\pi} \ge \frac{1}{2\pi} \|\psi(\alpha t, \omega/\alpha) - e^{i\omega t}\|_\infty \qquad \blacksquare$$

EXAMPLE 3.5. Franklin wavelets.

As was introduced in chapter 2, the Fourier transform of the scaling function of the Franklin wavelets is

$$\hat{\phi}(\omega) = \frac{\hat{\eta}(\omega)}{\sqrt{\sum_k \hat{\eta}^2(\omega + 2\pi k)}} = \frac{\sin^2\omega/2}{(\omega/2)^2}\left(1 - \frac{2}{3}\sin^2\frac{\omega}{2}\right)^{-\frac{1}{2}}$$

Thus

$$\hat{\phi}^*(\omega) = \sum_k \phi(k) e^{-i\omega n} = \sum_k \hat{\phi}(\omega + 2\pi k) = \left(1 - \frac{2}{3}\sin^2\frac{\omega}{2}\right)^{-\frac{1}{2}}$$

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^{*}(\omega)=\frac{\sin^2\frac{\omega}{2}}{(\omega/2)^2}$$

whose inverse Fourier transform is the second order spline function

$$S(t)=(1-|t|)\chi_{(-1,1)}$$

$$\psi(t,\omega)=\sum_{n=-\infty}^{+\infty} e^{i\omega t}S(t-n)=e^{i\omega[t]}S(t-[t])+e^{i\omega([t+1])}S(t-[t+1])$$

So, for $f\in U_\alpha$

$$f(t)=f(n/\alpha)(1+n-\alpha t)+f((n+1)/\alpha)(\alpha t-n) \quad \textit{when } n/\alpha\leq t<(n+1)/\alpha$$

Notice that S(t-[t])+S(t-[t+1])=1, so

$$|\psi(t,\omega)|\leq S(t-[t])+S(t-[t+1])=1$$

added the fact that $\psi(t,\omega)$ is of bounded variation on $[-\pi,\pi]$ gives

$$\|f-g\|_\infty\leq C\cdot \bigvee_{-\infty<\omega<\infty} \beta(\omega)$$

where C can be chosen as

$$C=\frac{1}{2\pi}\|\psi(t,\omega)-e^{i\omega t}\|_\infty=\frac{1}{\pi}$$

EXAMPLE 3.6. Quadratic spline.

Haar and Franklin wavelets correspond to first and second order basic splines. Here η is the basic spline of the third order with Fourier transform

$$\hat{\eta}(\omega)=[\frac{1-e^{-i\omega}}{i\omega}]^3$$

The $\{\eta(t-k)\}$ are not orthonormal. $\hat{S}=\hat{\eta}/\hat{\eta}^*$.

Since $\eta(1)=\eta(2)=1/2$, and $\eta(k)=0$, $k\neq 1,2$. Hence

$$\hat{\eta}^*(\omega)=\frac{1}{2}e^{-i\omega}(1+e^{-i\omega})$$

has a zero at $\omega=\pm\pi$. Therefore there is no sampling theorem in V_0.

EXAMPLE 3.7. Lemarié-Meyer Wavelet.

The next chapter will be devoted to the analysis of sampling theory regarding bandlimited scaling functions. Variations of the Lemarié-Meyer scaling function will also be discussed in great detail. It can be easily seen from the analysis of the next chapter that the sampling theorems exist for a special Lemarié-Meyer scaling function if and only if

$$\hat{\phi}^*(\omega)\neq 0.$$

4 Bandlimited Scaling function

The sampling theory and the error analysis technique of the last chapter are especially effective when the scaling function is bandlimited. As mentioned in Chapter 2, the Whittaker-Shannon sampling kernel is a scaling function. Clearly, it is bandlimited. Another family of bandlimited scaling functions is the bandlimited sampling scaling functions. Later, scaling functions with raised cosine spectrum will be presented. This is followed by the problem of intersymbol interference in communication.

4.1. Bandlimited Sampling Scaling Function

As was being mentioned in Chapter 2, the bandlimited sampling scaling function is formulated as follows.

Choose any $\hat{\eta} \in C^N(R)$, N being any positive integer and

(a) supp $\hat{\eta}=(-\pi-\varepsilon,\pi+\varepsilon)$, $0<\varepsilon\leq\pi$;

(b) $\hat{\eta}(\omega)=1$ for $|\omega|\leq\pi-\varepsilon$;

(c) $A\leq\Sigma_k|\hat{\eta}(\omega+2\pi k)|^2\leq B$, $0<A<1<B<\infty$;

(d) $\hat{\eta}(\omega)\geq 0$; and

(e) $\hat{\eta}(\omega)>0$, for $|\omega|<(\pi+\varepsilon)/2$

Introduce ϕ whose Fourier transform is given by

$$\hat{\phi}(\omega) = \frac{\hat{\eta}(\omega)}{(\sum_k |\hat{\eta}(\omega+2\pi k)|^2)^{1/2}}$$

then ϕ is an o.n. scaling function.

As was mentioned in chapter 2, this type of scaling function is a variation of Lemarié-Meyer scaling function. Henceforth, $\hat{\eta}$ and $\hat{\phi}$ will be defined as above. The asymptotic decay of ϕ is described below.

THEOREM 4.1. $\phi(x)=O(|x|^{-N-1})$ as $x \to \pm\infty$.

Proof. Use the following N-fold integration by parts

$$\int_{-\infty}^{\infty} \hat{\phi}^{(N)}(\omega) e^{ix\omega} d\omega = (-ix)^{-N} \int_{-\infty}^{\infty} \hat{\phi}(\omega) e^{ix\omega} d\omega = 2\pi(-ix)^N \phi(x)$$

First, if $\hat{\phi}^{(N+1)}$ exists and is in $L^1(R)$, then another integration by parts gives

$$|\phi(x)| \le C \cdot \frac{1}{|x|^{N+1}}$$

as $x \to \pm\infty$, and C is some positive constant. In general, since $\hat{\phi}^{(N)}$ is in L^1, for any $\varepsilon > 0$, function g can be found such that $g, g' \in L^1$ and $\| \hat{\phi}^{(N)} - g \|_1 < \varepsilon$ (see p25, Chui[1992a]). Let

$$f(x) = 2\pi(-ix)^N \phi(x)$$

$$h(x)=\int_{-\infty}^{\infty} g(\omega)\, e^{ix\omega} d\omega$$

then

$$|f(x)| \le |f(x)-h(x)| + |h(x)|$$

$$\le \|\hat{\phi}^{(N)}-g\|_1 + |h(x)| < \varepsilon + C\left|\frac{1}{x}\right|$$

to any fixed x, it follows that $f(x)=O(|x|^{-1})$ as $x\to\pm\infty$. So $\phi(x)=O(|x|^{-N-1})$ as $x\to\pm\infty$. ∎

In order to better understand the term $\hat{\phi}^*(\omega)$, the following theorem is established.

THEOREM 4.2. The Bandlimited-sampling scaling function ϕ satisfies

$$\hat{\phi}^*(\omega)=\sum_k \phi(k)\, e^{-i\omega k}=\sum_k \hat{\phi}(\omega+2\pi k)$$

and the last summation has at most two terms.

Proof. Since supp $\hat{\eta}=[-\pi-\varepsilon,\pi+\varepsilon]$, the support of derivative of $\hat{\phi}$, $\text{supp}(\hat{\phi})'\subseteq[-\pi-\varepsilon,\pi+\varepsilon]$. $(\hat{\phi})'$ is continuous, hence integrable on $[-\pi-\varepsilon,\pi+\varepsilon]$. By theorem B7 (see appendix B), $\hat{\phi}$ is of bounded variation. Since ϕ is bandlimited, $\Sigma_k\hat{\phi}(\omega+2\pi k)$ is a finite sum for any fixed ω, hence it's also continuous and of bounded variation on $[0,2\pi]$. By theorem B10 (see appendix B),

$$\sum_k \phi(k)\, e^{-i\omega k} = \sum_k \hat{\phi}(\omega+2\pi k)$$

everywhere. Because the width of supp$\hat{\phi}=2\pi+2\varepsilon<4\pi$, therefore

$$\sum_k \hat{\phi}(\omega+2\pi k)$$

has at most two terms. ▮

From condition (d), (e) in the definition of ϕ,

$$[\sum_k \hat{\eta}(\omega+2\pi k)]^2 \geq (\sum_k |\hat{\eta}(\omega+2\pi k)|^2) \geq A$$

Hence

$$\hat{\phi}^*(\omega) = \sum_k \hat{\phi}(\omega+2\pi k) = \frac{\sum_k \hat{\eta}(\omega+2\pi k)}{(\sum_k |\hat{\eta}(\omega+2\pi k)|^2)^{1/2}} > 0.$$

Using the above results, the scaling function ϕ satisfies all the conditions in (i), (ii) of theorem 3.1., theorem 3.2. and their related corollaries. More specifically, $\{\phi(n)\}\in l^1$ (because $\Sigma_n|\phi(n)|\leq\Sigma_n O(n^{-2})<\infty$), $\hat{\phi}(w)$ is bounded and bandlimited, $\phi(t)=O(x^{-1-\varepsilon})$ as $x\to\pm\infty$, $\hat{\phi}(w)\in L^1$ and $\hat{\phi}^*(w)\neq 0$.

Since $\hat{\eta}^*(\omega)=\Sigma_k\hat{\eta}(\omega+2\pi k)$ has at most two terms and $\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)=\hat{\eta}(\omega)/\hat{\eta}^*(\omega)$, Unlike the function $\hat{\phi}(\omega)$, $\hat{S}(\omega)$ is very easy to calculate, especially $\hat{S}(\omega)=\hat{\eta}(\omega)=\hat{\eta}^*(\omega)=1$ for $|\omega|\leq\pi-\varepsilon$.

Apply the theorems 3.1. and 3.2., the sampling theorem in the environment of bandlimited-sampling scaling function is summarized as follows.

Theorem 4.3. If $\hat{\phi}$ is a bandlimited-sampling scaling function, for $f \in V_0$, if $f(n)=\sum_k a_k \phi(n-k)$, then

$$f(t)=\sum_n f(n)S(t-n)$$

in the sense of $L^2(R)$, that is

$$\int_R |f(t)-\sum_{n=-N}^{N} f(n)S(t-n)|^2 dt \to 0 \text{ as } N\to\infty.$$

or if f(t) is continuous and is in $L^1(R)$, then

$$f(t)=\sum_n f(n)S(t-n)$$

where the convergence is both in $L^2(R)$ and $L^\infty(R)$ (uniform convergence), where

$$\hat{S}(\omega)=\hat{\phi}(\omega)/\hat{\phi}^*(\omega)$$

Using the same terminology as in chapter 3, S is called the sampling kernel. Apply theorem 3.3., the U_α version of the theorem is formulated as follows.

Corollary. Suppose ϕ is a bandlimited-sampling scaling function, for $f\in U_\alpha$, if $f(n/\alpha)=\Sigma_k a_k\phi(n-k)$, then

$$f(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\,S(\alpha t-n)$$

in the sense $L^2(R)$. Or if f(t) is continuous and is in $L^1(R)$, then

$$f(t)-\sum_{n=-\infty}^{\infty} f(n/\alpha)\,S(\alpha t-n)$$

and the convergence is both in $L^2(R)$ and $L^\infty(R)$, where S is the sampling kernel.

Notice that all the functions in $V_0=U_1$ are bandlimited to $[-\pi-\varepsilon,\pi+\varepsilon]$ and $f(x)\in U_\alpha \Leftrightarrow f(x/\alpha)\in U_1$. Functions in U_α are bandlimited to $[-\alpha(\pi+\varepsilon),\alpha(\pi+\varepsilon)]$. Hence S makes a very good low pass filter.

The error analysis in the last Chapter relies heavily on the property of the function $\Psi(t,\omega)=\Sigma_k S(t-n)e^{i\omega n}$.

Theorem 4.4. If $\hat{S}\in L^1(R)$, to every fixed t, the Fourier series of the function $\Phi(\omega)=\Sigma_k\hat{S}(\omega+2\pi k)e^{i(\omega+2\pi k)t}$ is $\Psi(t,\omega)=\Sigma_k S(t-n)e^{i\omega n}$. If $\Phi(\omega)$ is of bounded variation on $[0,2\pi]$, then $\Psi(t,\omega)$ equals to $(\Phi(\omega^+)+\Phi(\omega^-))/2$ everywhere.

Proof. Let g(x)=S(t+x) ,then

$$\hat{g}(\omega)=\int_{-\infty}^{\infty}S(t+x)\,e^{-i\omega x}dx=\int_{-\infty}^{\infty}S(x)\,e^{-i\omega x}\cdot e^{i\omega t}dx=e^{i\omega t}\hat{S}(\omega)$$

So $\Phi(\omega)=\Sigma_k\hat{g}(\omega+2\pi k)$ and $g\in L^1(R)$. Using theorem B11, the Fourier series of $\Phi(\omega)$ is

$$\sum_{n=-\infty}^{\infty}g(-n)\,e^{i\omega n}=\sum_n S(t-n)\,e^{i\omega n}$$

and the last statement is just a simple application of theorem B9. ∎

Like in the last chapter, when space U_α is considered, the sampling series g of f is defined as

$$g(t)=\sum_n f(n/\alpha)\,S(\alpha t-n)$$

Applying theorem 3.4., the following result is easily obtained.

Theorem 4.5. If ϕ is a bandlimited sampling scaling function and S is its relating sampling kernel. If $f\in L^2$ can be written as $f=f_0+h$, where $f_0\in V_0$, f_0 continuous, $f_0\in L^1$ and

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega t}d\beta(\omega)$$

where $\beta(\omega)$ is of bounded variation, then the sampling series g of f converges and

$$\|f-g\|_\infty \le C \cdot \bigvee_{-\infty\le\omega\le+\infty} \beta(\omega)$$

where C is a constant.

Proof. If the partial sum of

$$\psi(t,\omega) = \sum_{n=-\infty}^{+\infty} e^{i\omega n} S(t-n)$$

is proved to be uniformly bounded for all t and ω, then the result of the theorem is just a simple application of theorem 3.4.

Since $\hat{\phi}$ is of bounded variation (see the proof of theorem 4.2.) and $|\hat{\phi}^*(w)|$ is bounded away from zero (see theorem 3.1. or 3.2), suppose $|\hat{\phi}^*(w)| \ge m$, then $e^{-i\omega t}/\hat{\phi}^*(w)$ is bounded, by theorem B8,

$$\hat{S}(\omega)\, e^{-i\omega t} = \frac{\hat{\phi}(\omega)}{\hat{\phi}^*(\omega)}\, e^{-i\omega t}$$

is of bounded variation. Since the summation $\Phi(\omega)=\sum_k \hat{S}(\omega+2\pi k)e^{i(\omega+2\pi k)t}$ is a finite sum, it is also of bounded variation and continuous. By theorem 4.4., we have $\psi(t,\omega)=\Phi(\omega)$ and by theorem B9 the partial sum of $\psi(t,\omega)$ is uniformly bounded.

Since $|\hat{S}|\leq 1/m$ and $\Phi(\omega)$ has at most two terms, the constant C can be chosen

$$C=\frac{1}{2\pi}\|\psi(t,\omega)-e^{i\omega t}\|_\infty\leq\frac{1}{2\pi}\left(\frac{2}{m}+1\right) \qquad \blacksquare$$

The U_α version of the theorem is formulated as follows.

Corollary. If $\hat{\phi}$ is a bandlimited sampling scaling function and S is its relating sampling kernel. If $f\in L^2$ can be written as $f=f_\alpha+h$, where $f_\alpha\in U_\alpha$, f_α continuous, $f_\alpha\in L^1$ and

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega t}d\beta(\omega)$$

where $\beta(\omega)$ is of bounded variation, then the sampling series g of f converges and

$$\|f-g\|_\infty\leq C\cdot\bigvee_{-\infty\leq\omega\leq+\infty}\beta(\omega)$$

where C is a constant.

4.2. Scaling functions with Raised-Cosine Spectrum

Let $0<\gamma\leq 1$, and define

$$\hat{\eta}(\omega)=\begin{cases} 1 & 0\leq\omega\leq\pi(1-\gamma) \\ \frac{1}{2}\{1-\sin[\frac{1}{2\gamma}(\omega-\pi)]\} & \pi(1-\gamma)\leq\omega\leq\pi(1+\gamma) \\ 0 & \omega\geq\pi(1+\gamma) \\ \hat{\eta}(-\omega) & \omega<0 \end{cases}$$

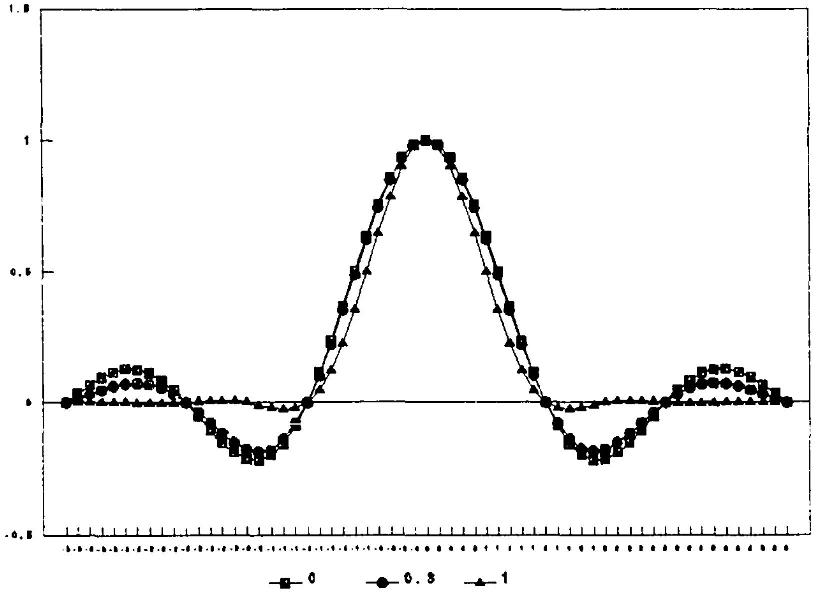
then

$$\hat{\eta}'(\omega)=\begin{cases} 0 & 0\leq\omega\leq\pi(1-\gamma) \\ \frac{1}{4\gamma}\cos[\frac{1}{2\gamma}(\omega-\pi)] & \pi(1-\gamma)\leq\omega\leq\pi(1+\gamma) \\ 0 & \omega\geq\pi(1+\gamma) \\ \hat{\eta}(-\omega) & \omega<0 \end{cases}$$

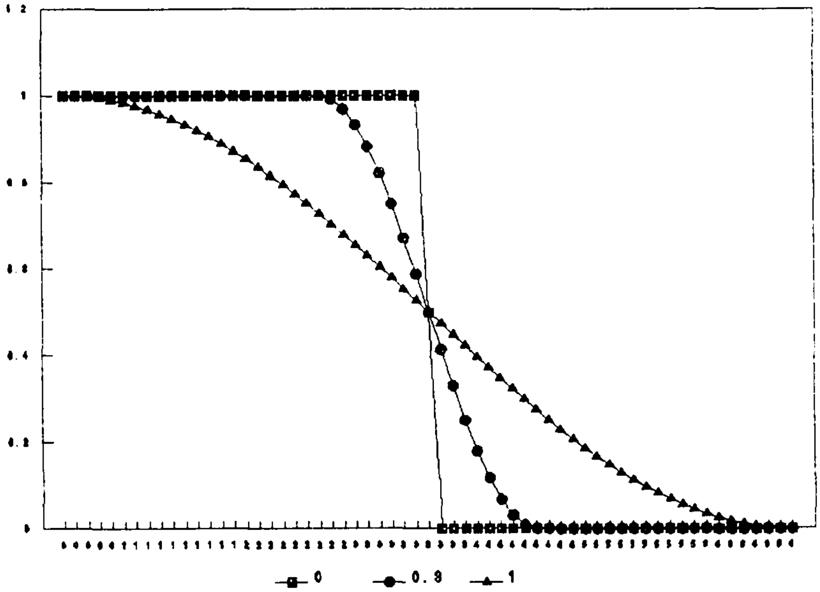
So $\hat{\eta}\epsilon C^1(R)$ and satisfy all the other conditions to make ϕ a bandlimited sampling scaling function.

Notice that $\Sigma_k\hat{\eta}(\omega+2\pi k)=1$ everywhere, hence $\hat{S}=\hat{\eta}/\Sigma_k\hat{\eta}(\omega+2\pi k)=\hat{\eta}$ and

$$S(t)=\eta(t)=\frac{\sin\pi t}{\pi t}\frac{\cos\gamma\pi t}{1-4\gamma^2t^2}$$



Time response when γ=0, 0.3, 1.



Frequency response when γ=0, 0.3, 1.

$S(t)=\eta(t)=O(|t|^{-3})$, as $t\to\pm\infty$. The limit function of S(t) as $\gamma\to 0$ is the Whittaker-Shannon sampling kernel. The Whittaker-Shannon sampling kernel can be viewed as a function with raised-cosine spectrum where $\gamma=0$. The sampling theorem for functions with raised-cosine spectrum is just a restatement of the corollary of theorem 4.3.

THEOREM 4.6. Suppose ϕ is a scaling function derived from the raised cosine spectrum. For $f\in U_\alpha$, if $f(n/\alpha)=\Sigma_k a_k\phi(n-k)$, then

$$f(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\,S(\alpha t-n)$$

in the sense $L^2(R)$. Or if f(t) is continuous and is in $L^1(R)$, then

$$f(t)=\sum_{n=-\infty}^{\infty} f(n/\alpha)\,S(\alpha t-n)$$

and the convergence is both in $L^2(R)$ and $L^\infty(R)$, where

$$S(t)=\eta(t)=\frac{\sin\pi t}{\pi t}\,\frac{\cos\gamma\pi t}{1-4\gamma^2t^2}$$

The aliasing error analysis is also just a simple application of the corollary of theorem 4.5.

THEOREM 4.7. Let ϕ be a scaling function with raised cosine spectrum and $S=\phi$ is its relating sampling kernel. If $f\in L^2$ can be written as $f=f_\alpha+h$, where $f_\alpha \in U_\alpha$, f_α continuous, $f_\alpha \in L^1$ and

$$h(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega t} d\beta(\omega)$$

where $\beta(\omega)$ is of bounded variation, then the sampling series g of f converges and

$$\|f-g\|_\infty \le C \cdot \bigvee_{-\infty\le\omega\le+\infty} \beta(\omega)$$

where C can be chosen as

$$C=\frac{3}{2\pi}$$

Proof. From the proof of theorem 4.5. it can be noticed that since $|S|\le 1$, $|\Psi(t,\omega)|\le 2$. Therefore

$$C=\frac{1}{2\pi}(\|\psi(t,\omega)\|_\infty+1)\le\frac{3}{2\pi} \qquad \blacksquare$$

The spectral version of the sampling series

$$\hat{g}(\omega)=\hat{S}(\omega/\alpha)\sum_n f(n/\alpha)\, e^{-i\omega n/\alpha}=\hat{S}(\omega/\alpha)\sum_n \hat{f}(\omega+2\alpha\pi n)$$

is obtained by applying the Fourier transform to both sides of the sampling series

$$g(t)=\sum_n f(n/\alpha)\, S(\alpha t-n)$$

Since $\hat{S}(\omega)$ is of finite support, using the sampling series acts like a low pass filter. When supp $\hat{f}\subseteq[\alpha(-\pi+\varepsilon),\alpha(\pi-\varepsilon)]$

$$\sum_{n}\hat{f}(\omega+2\alpha\pi n)=\hat{f}(\omega)$$

and $\hat{S}(\omega/\alpha)=1$ for $[\alpha(-\pi+\varepsilon),\alpha(\pi-\varepsilon)]$. If the sampling theorem in U_α is used, then

$$\hat{g}(\omega)=\hat{S}(\omega/\alpha)\hat{f}(\omega)=\hat{f}(\omega)$$

hence using sampling theorem means perfect reconstruction of the signals. If $f\notin U_\alpha$, then error bounds can also be obtained by using theorem 4.7. Clearly, the spaces U_α corresponds to different bandwidth filters.

The advantage of using the raised-cosine spectrum instead of the Whittaker-Shannon sampling kernel is the continuity in the spectral domain. The spectrum of Whittaker-Shannon sampling kernel is not continuous. The fall off rate are $O(t^{-3})$ for raised-cosine spectrum and $O(t^{-1})$ for Whittaker-Shannon sampling kernel. This makes the raised-cosine spectrum better than Whittaker-Shannon sampling kernel for certain applications. The comparison will be more clear after the issue of intersymbol interference is described.

4.3. Intersymbol Interference.

The function of a typical digital communication system can be abstracted into the following 5 parts.

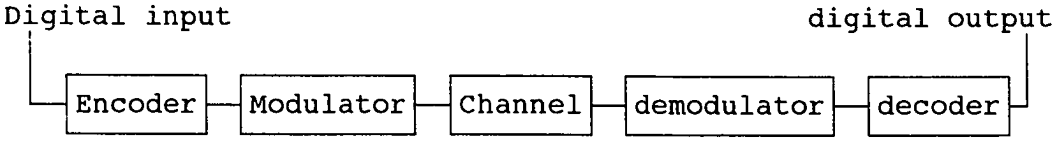


Fig 4.3. Block diagram of a digital communication system

The **encoder** converts the input digital sequence into another digital sequence for the purpose of error detection, error correction and/or encryption. For example, adding a parity checking bit can detect an error occurring in the output.

The **modulator** operates on the output of the encoder and produce a signal suitable for transmission. In radio or television broadcasting, certain stations are allowed to transmit signals only within a certain bandwidth, so the modulator has to produce such a signal.

The **channel** is the medium used to transmit the signal. It can be twisted pairs, coaxial cables, optical cables, or microwave. During transmission, the signal may be affected by noise. Noise can be a small fluctuation of air waves or lightening bursts. Smart modulator-demodulator schemes are designed to combat noise situations.

The **demodulator** operates on the received signal and tries to reproduce the original message transmitted. Theoretically, it performs the mathematical inverse of the modulator.

The **decoder** performs the inverse function of the encoder. It decrypts and performs error detection/correction.

In radio and television broadcasting, each station is assigned a certain band in the frequency domain in order to avoid signal collision from different stations. This means, after modulation, the signal which is about to be transmitted has to be bandlimited. By an elementary property of the Fourier transform

$$\text{If } \mathcal{F}[f(t)] = F(\omega), \text{ then}$$

$$\mathcal{F}[e^{iat}f(t)] = F(\omega - a) \text{ and } \mathcal{F}[f(t/a)] = F(a\omega)$$

If the problem of generating a function bandlimited to $[-\sigma,\sigma]$ is solved, by a simple change of phase and scale, resulting in a function bandlimited to any finite interval.

In communications, the process of generating a function bandlimited to $[-\sigma,\sigma]$ at the output of the modulator is called **baseband pulse shaping.**

Baseband pulse shaping is accomplished by a certain bandlimited function g(t), called the **basic pulse shape.** To any digital sequence $\{a_n\}$, the function formed at the output of the modulator is

$$\sum_n a_n g(t-n) .$$

Another criteria in choosing the basic pulse shape g(t) depends on how $\{a_n\}$ can be recovered accurately and effectively after receiving the signal and passing the demodulator.

In communications, if the signal distortion is due to the overlap in time between successive symbols then the error is known as **intersymbol interference.** In the ideal channel, the receiving signal is equal to the transmitted signal. That is that the received signal is

$$y(t)=\sum_n a_n g(t-n) .$$

Engineers are most interested in the received signal at the sample instances, that is

$$y(k)=\sum_n a_n g(k-n) .$$

In order to avoid intersymbol interference, g(n) has to be zero for all integer $n \neq 0$. Using appropriate normalization, g(t) has to satisfy

$$g(n) = \delta_{0n}$$

The Whittaker-Shannon sampling kernel clearly satisfies this criteria. By using it as the basic pulse shape will result in no intersymbol interference. The same thing happens when the function with raised cosine spectrum is employed since it also satisfies $S(n)=\delta_{0n}$.

Other factors also determine the choice of the basic pulse shape. Usually the timing for sampling is very hard to determine, so the fall off rate of the basic pulse shape is preferred to be large. More precisely, since $y(k+\Delta k)=A\Sigma_n a_n g(k+\Delta k-n)+\xi(k+\Delta k)$, the term $\Sigma_{n\neq 0} a_n g(k+\Delta k-n)$ is minimized. When g(t) is the Whittaker-Shannon sampling kernel, an extreme case occurs using $a_n=(-1)^n$, here $\Sigma_n a_n g(1/2-n)$ is not convergent. Hence a function g such that $\Sigma_n |g(t-n)|$ converge for all t is preferred. From this point of view, the functions with raised-cosine spectrum seem to be a much better choice.

Another consideration is that actually Whittaker-Shannon sampling kernel and the functions with raised-cosine spectrum are all not realizable in finite time. Recall a bandlimited function other than zero can not be time limited. Hence the fall off rate of the basic pulse shape is preferred to be large so that it can closely approximate real world time limited signals. From this viewpoint, functions with raised-cosine spectrum are again preferred.

5 Conclusion

In this dissertation, solutions to the problem of creating sampling theorems in wavelets subspaces is given. Several sufficient conditions are presented for sampling theorems to hold (theorems 4.1., 4.2. and corollaries). The notion of continuous multiresolution analysis is introduced and sampling theorems are extended to different (scale) wavelet subspaces (theorem 3.3. and 3.4). Aliasing error analysis techniques based on the idea of "band covering" is also developed (theorem 3.5.).

The sampling theory in the wavelet subspace created by a special class of scaling functions, namely bandlimited sampling scaling functions, is also proven and are easy to apply (theorem 4.1.-4.7.). A special example of bandlimited sampling scaling function, namely those with raised cosine spectrum is introduced along with its sampling theory. The connection between the sampling theory and the problem of intersymbol interference is also presented.

Appendix A. Banach spaces and Hilbert Spaces.

Concepts and theorems in the appendices will be given without or with little proof.

A.1. Banach Space.

A **vector space** is an abstract space which has an algebraic structure. Formally, a set X of elements is called a vector space over the field of complex number C (or over some other field), if there is a function + defined from X×X to X and a function · defined from C×X to X such that

i. Associative law holds: $(x+y)+z=x+(y+z)$.

ii. Commutative law holds: $x+y=y+x$.

iii. Existence of zero: there exists θ s.t. $x+\theta=x$ for all $x\in X$

iv. $a(x+y)=ax+ay$ for all $a\in C$, $x,y\in X$.

v. $(a+b)x=ax+bx$ for all $a,b\in C$, $x\in X$.

vi. $a(b\cdot x)=(ab)\cdot x$ for all $a,b\in C$, $x\in X$.

vii. $0\cdot x=\theta$, $1\cdot x=x$.

Only vector spaces over the complex field will be used in the following. A nonnegative real-real valued function $\| \ \|$ defined on a vector space X is called a **norm** if

i. $\|x\|=0 \Leftrightarrow x=\theta$, where θ is the zero of the vector space.

ii. $\|x+y\| \le \|x\|+\|y\|$ for all $x,y \in X$.

iii. $\|ax\|=a\|x\|$, for all $a \in C$, $x \in X$.

A normed space becomes a **metric space** if the metric η is defined by $\eta(x,y)=\|x-y\|$. If under this metric the vector space is complete, it is called a **Banach space.**

Example A1: For each p, $1 \le p < \infty$, let $L^p(R)$ denote the class of measurable functions f such that the Lebesgue integral

$$\int_{-\infty}^{+\infty} |f(x)|^p dx$$

is finite. Also, let $L^\infty(R)$ be the class of functions bounded almost everywhere (a.e.). Choose $\|f\|_p$ to be

$$\|f\|_p = \begin{cases} \{\int_{-\infty}^{\infty} |f(x)|^p dx\}^{1/p} & for\ 1 \le p < \infty \\ \underset{-\infty < x < \infty}{ess\,sup} |f(x)| & for\ p = \infty \end{cases}$$

then $\|\cdot\|_p$ satisfy the conditions (ii) and (iii) of the norm axiom. From $\|f\|_p=0$, it follows that f=0 a.e. However, two measurable function are considered to be equivalent if they are equal a.e. and functions within an equivalence class are not distinguished. Every $L^p(R)$, $1 \le p \le \infty$ is a Banach space.

Similarly, the space $L^p[a,b]$ and l^p can be formulated as follows.

Example A2: For each p, $1 \le p < \infty$, $L^p[a,b]$ denote the class of measurable functions f such that the Lebesgue integral

$$\int_a^b |f(x)|^p dx$$

is finite. Also, let $L^\infty[a,b]$ be the class of functions bounded almost everywhere on [a,b]. The norm of $L^p[a,b]$ is defined as

$$\|f\|_p = \begin{cases} \{\int_a^b |f(x)|^p dx\}^{1/p} & \text{for } 1 \le p < \infty \\ \operatorname{ess\,sup}_{a<x<b} |f(x)| & \text{for } p = \infty \end{cases}$$

each $L^p[a,b]$, $1 \le p \le \infty$ is a Banach space. The reason is similar to $L^p(R)$, the elements of $L^p[a,b]$ are not functions but rather equivalence classes of functions.

Example A3: For each p, $1 \le p < \infty$, l^p denote the Banach space consisting of all infinite sequences $c = \{c_1, c_2, c_3, \ldots\}$ such that $\|c\|_p = (\sum_{n=1}^\infty |c_n|^p)^{1/p} < \infty$ and l^∞ denote the space of all infinite sequences which are bounded, i.e. $\|c\|_\infty = \sup_n\{|c_n|\} < \infty$.

A.2. Basis in Banach Space.

The most widely used basis in a Banach space is the Schauder basis (Schauder [1927]).

Definition A1. A sequence of vectors $\{x_1, x_2, x_3, \ldots\}$ in an infinite-dimensional Banach space X is said to be a **Schauder**

basis for X if to each vector x in the space there corresponds a unique sequence of scalars $\{c_1, c_2, c_3, \ldots\}$ such that

$$x = \sum_{n=1}^{\infty} c_n x_n .$$

The convergence of the series is with respect to the strong (norm) topology of X; in other words,

$$\| x - \sum_{i=1}^{n} c_i x_i \| \to 0 \qquad as\ n \to \infty .$$

Henceforth, the term basis for an infinite-dimensional Banach space will always mean a Schauder basis.

Example A4. The Banach space $l^p (1 \le p < \infty)$ consists, by definition, of all infinite sequences of scalars $c = \{c_1, c_2, c_3, \ldots\}$ such that $\| c \|_p = (\sum_{n=1}^{\infty} |c_n|^p)^{1/p} < \infty$. The vector operations are coordinatewise. In each of these spaces, the "natural basis" $\{e_1, e_2, e_3, \ldots\}$, where

$$e_n = \{0, 0, \ldots 0, 1, 0, \ldots\}$$

and the 1 appears in the nth position, is easily seen to be a Schauder basis.

Example A5: The Banach space C[a,b] consists of all the continuous functions on the closed finite interval [a,b] together with the norm $\|f\|=\max|f(x)|$. Let $\{x_0,x_1,x_2,...\}$ be countable dense subset of $[a,b]$ with $x_0=a$ and $x_1=b$. Set

$$f_0=1 \quad \text{and} \quad f_1=(x-a)/(b-a).$$

When n>1, the set of points $\{x_0,x_1,...,x_{n-1}\}$ partition $[a,b]$ into disjoint open intervals, one of which contains x_n; call it I. Define

$$f_n=\begin{cases} 0 & \text{if } x\notin I \\ 1 & \text{if } x=x_n \\ \text{linear} & \text{otherwise} \end{cases}$$

for n=2,3,4....The sequence $\{f_0,f_1,f_2,...\}$ will be a basis for C[a,b].

If a vector space has a countable basis, it is called **separable.** A Banach space with a basis must be separable. Because if $\{x_n\}$ is a basis for X, the set of all finite linear combinations $\Sigma c_n x_n$ where c_n are rational, is countable and dense in X. It follows, for example, l^∞ is not separable, it cannot possess a basis.

Not every separable Banach space has a basis. This was proved by Per Enflo[1973], who constructed an example of a separable Banach space having no basis. Some familiar Banach spaces: C[a,b], l^p, L^p. $(1\le p<\infty)$ are examples of separable Banach spaces.

Among the most important Banach spaces are the spaces L^p $(1 \le p \le \infty)$. Among most famous and important inequalities in the L^p norm are the Holder inequality and Minkowski inequality.

Theorem A1 (Holder inequality). If $1 \le p, q \le \infty$, $1/p+1/q=1$ (when one of them is ∞, the other is understood to be 1) and if $f \in L^p$, $g \in L^q$, then $fg \in L^1$ and

$$\int_E fg \le \|f\|_p \cdot \|g\|_q.$$

(see p113, Royden [1968]).

Theorem A2 (Minkowski Inequality). If $f, g \in L^p$, then $f+g \in L^p$ and

$$\|f+g\|_p \le \|f\|_p + \|g\|_p$$

(see p114, Royden [1968]).

Minkowski integral inequality is as follows.

Theorem A3 (Minkowski integral Inequality). If

$$\int_{E_1}\int_{E_2} f(x,y) \, dxdy < \infty$$

then

$$\{\int_{E_1} \int_{E_2} f(x,y) \, dx \;\; ^p dy\}^{1/p} \le \int_{E_2} \{\int_{E_1} f(x,y) \;\; ^p dy\}^{1/p} dx$$

(see p3, Stein [1971]).

Among the most widely used operation in electrical engineering is **convolution** defined as

$$h(y)=\int_E f(y-x)\,g(x)\,dx$$

and denoted as h=f*g. The following is an important property of convolution.

Theorem A4. If $f\in L^p$ $(1\le p\le\infty)$ and $g\in L^1$ then h=f*g is well defined and belongs to L^p. Moreover

$$\|h\|_p\le\|f\|_p\|g\|_1 .$$

Proof. By Minkowski integral inequality

$$(\int_E h(x)\quad {}^p dx)^{1/p}=\{\int_E \int_E f(y-x)\,g(y)\,dx\quad {}^p dy\}^{1/p}$$

$$\le\int_E\{\int_E f(y-x)\,g(x)\quad {}^p dy\}^{1/p}dx=\int_E\{\int_E f(y-x)\quad {}^p dy\}^{1/p}\ g(x)\ dx$$

$$=\|f\|_p\|g\|_1$$ ∎

A.3. Hilbert Space

A **Hilbert space** is a Banach space H in which there is a function (x,y) on H×H to C with the following properties.

i. $(ax_1+bx_2,y)=a(x_1,y)+b(x_2,y)$

ii. $(x,y)=\overline{(y,x)}$

iii. $(x,x)=\|x\|^2$

(x,y) is called the **inner product** of x and y.

Example A6. The space R^n with

$$(x,y)=\sum_{i=1}^{n} x_i y_i .$$

is a Hilbert space.

Example A7. Space L^2 with inner product

$$(x,y)=\int_E x(t)\overline{y(t)}\,dt.$$

is a Hilbert space.

Among the most well known inequalities in a Hilbert space is the **Cauchy-Buniakowsky-Schwarz (CBS) inequality** (sometimes called Schwarz inequality or Cauchy-Schwarz inequality)

Theorem A5 (CBS inequality). If f and g belong to a Hilbert space, then

$$|(x,y)| \leq \|x\|\cdot\|y\|$$

where the equality can only occur when x and y are linearly dependent.

A.4. Basis in Hilbert Space.

Two element x,y in H are **orthogonal** if $(x,y)=0$. An set S in H is called an **orthogonal system** if any two different elements x and y of S are orthogonal, that is $(x,y)=0$. An orthogonal system S is called **orthonormal** if $\|x\|=1$ for any $x \in S$.

Theorem A6. Every separable Hilbert space has an orthonormal basis (p212, Royden [1968]).

The most important property of an orthonormal basis is the simplicity of the basis expansions. If $\{e_1, e_2, e_3, \ldots\}$ is an orthonormal basis for a Hilbert space H, then for every $f \in H$ the **Fourier expansion** is given by

$$f=\sum_{n=1}^{\infty} (f, e_n)\, e_n .$$

The inner product (f,e_n) is called the nth Fourier coefficient of f. And the following isometry holds.

Theorem A7 (Parseval's identity).

$$\|f\|^2=\sum_{n=1}^{\infty} |(f,e_n)|^2.$$

(see p212, Royden [1968]).

The validity of Parseval's identity for every vector in the space is both necessary and sufficient for an orthonormal sequence to be a basis.

Similarly, the **generalized Parseval identity** holds:

$$(f,g)=\sum_{n=1}^{\infty} (f,e_n)\overline{(g,e_n)}.$$

Even if an orthonormal sequence $\{e_n\}$ is incomplete, **Bessel's Inequality** is always valid:

$$\sum_{n=1}^{\infty} |(f,e_n)|^2 \leq \|f\|^2.$$

whenever $f \in H$. This shows, that the Fourier coefficients of each element of H form an l^2 sequence. Conversely, is the **Riesz-Fisher theorem,**

Theorem A8 (Riesz-Fisher). For every l^2 sequence $\{c_n\}$, there exists an $f \in H$ for which

$$(f,e_n)=c_n, \quad n=1,2,3,\ldots.$$

Proof. simply choose $f=\Sigma_n c_n e_n$. ▌

So, if $\{e_n\}$ is an o.n. basis of a separable Hilbert space H, then $f\rightarrow\{(f,e_n)\}$ is a Hilbert space isomorphism between H and l^2. It follows that from the geometric point of view all separable Hilbert space are "indistinguishable", that is, isomorphic.

Example A8. In l^2, the "natural basis" $\{e_1,e_2,e_3,...\}$, where

$$e_n=\{0,0,...0,1,0,...\}$$

and the 1 appears in the nth position is orthonormal.

Example A9. In $L^2[-\pi,\pi]$, with the inner product

$$(x,y)=\frac{1}{2\pi}\int_{-\pi}^{\pi}x(t)\overline{y(t)}dt.$$

the complex trigonometric system $\{e^{int}\}_{-\infty}^{\infty}$ constitutes an o.n. basis.

A.5. Linear Operators.

A mapping T of a vector X into a vector space Y is called a linear mapping, a linear operator, or a linear transformation if

$$T(ax_1+bx_2)=aT(x_1)+bT(x_2)$$

for all x_1, x_2 in X and all scaler a, b. If X and Y are normed vector spaces, a linear operator T is called **bounded** if to all $x \in X$:

$$\|Tx\|_Y \leq M\|x\|_X$$

for some M>0. The inf of such M is defined is the norm of T and denoted by $\|T\|$, hence

$$\|T\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Tx\|_Y}{\|x\|_X}$$

Clearly to any $x \in X$

$$\|Tx\|_Y \leq \|T\|\|x\|_X$$

The linear mapping have several important properties.

Theorem A9. If T is a linear mapping from X to Y, then the following three conditions are equivalent

i. T is bounded,

ii. T is continuous,

iii. T is continuous at one point.

(see p184, Royden [1986]).

Theorem A10. The space of all bounded linear mapping from a normed vector space X to a Banach space Y is itself a Banach space (see p185, Royden [1968]).

For any subset O of X, let T[O] denote the image of O under T in Y. The following important definition uses this notation.

Definition A2. A mapping T from X to Y is called **open mapping** if T[O] is open in Y whenever O is open in X.

Theorem A11 (Open mapping theorem). A bounded linear mapping T from a Banach space X onto a Banach space Y is an open mapping (see p195, Royden [1986]).

A important class of linear mappings is the linear functionals.

Definition A3. A **linear functional** of a vector space X is a linear operator from X to R. Here R denotes the vector space of all real numbers under the usual addition and multiplication.

The space of bounded linear functional on a normed space X is called the **dual** of X and is denoted by X^*. Since R is a Banach space, the dual X^* of any normed space X is a Banach space.

If g is a function in L^q $(1/p+1/q=1)$, by Holaer's inequality, the functional F on L^p

$$F(f)=\int_E fg$$

is a bounded linear mapping. More exactly, the following result holds.

Theorem A12. Each function g in L^q defines a bounded linear functional F on L^p by

$$F(f) = \int_E fg$$

and $\|F\| = \|g\|_q$ (see p119, Royden [119]).

The converse of the above theorem is also true for $1 \le p < \infty$, that is, every bounded linear functional is obtained in this manner.

Theorem A13 (Riesz representation Theorem). Let F be a bounded linear functional on L^p, $1 \le p < \infty$. Suppose q satisfies $1/p+1/q=1$, then there is a $g \in L^q$ such that

$$F(f) = \int_E fg$$

and $\|F\| = \|g\|_q$. (see p121, Royden [1968]).

A special case of the above theorem is when p=2. Here q=2, this means that any bounded linear functional of the Hilbert space L^2 can be represented as an inner product of the space. In fact, this is true for any Hilbert space. More exactly, bounded linear functional on the Hilbert space can be written as an inner product within the space.

Theorem A14. Let F be a bounded linear functional on the Hilbert space H. Then there is a $y \in H$ such that $F(x)=(x,y)$ for all x. Moreover, $\|f\|=\|y\|$. (p213, Royden[1968])

By the CBS inequality, $(x,y) \leq \|x\| \cdot \|y\|$. For any $y \in H$, (x,y) is a bounded linear functional on H. This makes $y \rightarrow (\cdot,y)$ an isomorphism between H and H^*.

A.6. Functional Hilbert Space.

Most of the important examples of Hilbert spaces are function space. The special properties of the functions considered, such as analyticity, enrich the structure of the space, which in turn supplies added information about the functions.

Let H be a Hilbert space whose elements are real or complex-valued functions defined on a set T. Call H a **functional Hilbert space** if all the evaluation functionals

$$f \rightarrow f(t)$$

$f \in H$, for each fixed $t \in T$ are continuous. By the Riesz representation theorem, for $t \in T$ there exists a unique element $K_t \in H$ such that

$$f(t) = \langle f, K_t \rangle \quad f \in H.$$

The function K on TxT, defined by

$$K(t,u)=(K_t,K_u)=K_t(u),$$

is called the **reproducing kernel** of H.

Remark. Not every Hilbert space is functional Hilbert space, e.g. L^2 is not. Some books called functional Hilbert space **reproducing kernel Hilbert space (RKHS).**

Example A10. The Paley-Wiener space PW_π is a functional Hilbert space with reproducing kernel

$$K(t,u)=\frac{\sin\pi(t-u)}{\pi(t-u)},$$

i.e., for $f\in PW_\pi$

$$f(t)=\int_{-\infty}^{\infty}K(t,u)\,f(u)\,du.$$

The reproducing kernel of the functional Hilbert space can always be described explicitly in terms of an orthonormal basis for the space.

Theorem A15. If $\{e_1,e_2,e_3,\ldots\}$ is an orthonormal basis for a functional Hilbert space with reproducing kernel K, then

$$K(t,u)=\sum_{n=1}^{\infty}\overline{e_n(t)}\,e_n(u).$$

Proof. By the generalized Parseval's identity

$$K(t,u)=(K_t,K_u)=\sum_{n=1}^{\infty}(K_t,e_n)\overline{(K_u,e_n)}=\sum_{n=1}^{\infty}\overline{e_n(t)}e_n(u). \quad \blacksquare$$

Example A11. In view of the above theorem, the reproducing kernel of the Paley-Wiener space PW_π is also equal to

$$K(t,u)=\sum_{n=-\infty}^{\infty}\frac{\sin\pi(t-n)}{\pi(t-n)}\frac{\sin\pi(u-n)}{\pi(u-n)}.$$

A.7. Biorthogonal sequence

If $\{x_1,x_2,x_3,\ldots\}$ is a basis for banach space X, then every vector x in X has a unique expansion

$$x=\sum_{n=1}^{\infty}c_n x_n.$$

Clearly each c_n is a linear function of x. If the linear functional is denoted by f_n, then $c_n=f_n(x)$ and

$$x=\sum_{n=1}^{\infty}f_n(x)x_n.$$

The functionals are called the **coefficient functionals** associated with the basis $\{x_n\}$. Clearly they all belong to X^*, the space of all bounded linear functionals on X (X^* is usually called the **Dual space** of X).

$\{f_n\}$ generally does not constitutes a basis of X^*. For if X^* is nonseparable, then it contains no basis at all, for example $X=l^1$. But $\{f_n\}$ does form a basis of its closed linear span. The closed linear span of a sequence $\{x_n\}$ will be denoted by $[x_n]$.

Theorem A16. Suppose that $\{x_n\}$ is basis of a Banach space X and let $\{f_n\}$ be the associated sequence of coefficient functionals. Then $\{f_n\}$ is a basis for $[f_n]$ and the expansion

$$f=\sum_{n=1}^{\infty} f(x_n)\, f_n$$

is valid for every f in $[f_n]$ (see p27, Young [1980]).

When X is a **reflective** Banach space (X^{**} isomorphic to X), $\{f_n\}$ is a fortiori complete.

Theorem A17. If $\{x_n\}$ is a basis for a reflexive Banach space, then the associated sequence of coefficient functionals $\{f_n\}$ is a basis for X^* (see p28, Young [1980]).

In any Hilbert space H, by the CBS inequality,

$(x,y) \le \|x\| \cdot \|y\|$. For any $y \in H$, (x,y) is a bounded linear functional on H. This makes $y \to (\cdot, y)$ an isomorphism between H and H^*. Therefore every Hilbert space is a reflective Banach space.

Two sequences $\{x_n\}$ and $\{y_n\}$ in H are said to be **biorthogonal** if

$$(x_m, y_n) = \delta_{mn}$$

for all m and n. It is clear an o.n. sequence is biorthogonal to itself.

Suppose $\{x_n\}$ is a basis for Hilbert space H. The isomorphism between H and H^* shows that to each coefficient functional f_n there corresponds an $y_n \in H$ such that $f_n(x) = (x, y_n)$ for all x. Since $f_n(x_m) = \delta_{mn}$, $\{x_n\}$ and $\{y_n\}$ are biorthogonal.

By theorem A17 and the fact every Hilbert space is reflective, $\{y_n\}$ is also a basis of H and it's called the **dual basis** of $\{x_n\}$. In terms of the representation, each vector in can be uniquely written in the form

$$x = \sum_n (x, y_n) x_n .$$

Since $\{y_n\}$ is also a basis of H, then

$$x=\sum_n (x,x_n)y_n.$$

A.8. Equivalent Bases.

Definition A4. Two bases $\{x_n\}$ and $\{y_n\}$ are **equivalent** if

$$\sum_{n=1}^{\infty} c_n x_n \text{ is convergent iff } \sum_{n=1}^{\infty} c_n y_n \text{ is convergent.}$$

The equivalent bases can be completely characterized by an isomorphism of the space.

Theorem A18. Two bases $\{x_n\}$ and $\{y_n\}$ are equivalent if and only if there exists a bounded invertible operator $T:X\to X$ such that $Tx_n=y_n$ for all n (see p30, Young [1980])

In a separable Hilbert space the most important bases are orthonormal. Second in line are those bases equivalent to some orthonormal basis. They are called Riesz basis.

Definition A5. A basis for a Hilbert space is a **Riesz basis** if it is equivalent to an orthonormal basis.

Theorem A19. In a Hilbert space equivalent bases have equivalent biorthogonal sequences. Therefore, the dual basis of a Riesz basis is a Riesz basis (see p31, Young[1980]).

The next theorem provides a number of important properties of a Riesz basis.

Theorem A20. Let H be a separable Hilbert space. Then the following statement are equivalent.

(1) The sequence $\{f_n\}$ forms a Riesz basis for H.

(2) There is an equivalent inner product on H (two inner product are equal if they generates equivalent norms), with respect to which $\{f_n\}$ becomes an orthonormal basis for H.

(3) The $\{f_n\}$ is complete in H, and there exists B≥A>0 such that for any l^2 sequence $\{c_k\}$ it follows that

$$A\sum_k |c_k|^2 \le \|\sum_k c_k f_i\|^2 \le B\sum_k |c_k|^2$$

(4) The sequence $\{f_n\}$ is complete in H, and its Gram matrix

$$((f_i, f_j))_{i,j-1}^{\infty}$$

generates a bounded invertible operator on l^2.

(5) The sequence $\{f_n\}$ is complete in H and possesses a complete biorthogonal sequence $\{g_n\}$ such that

$$\sum_{n-1}^{\infty} (f, f_n)^{2} < \infty \quad and \quad \sum_{n-1}^{\infty} (f, g_n)^{2} < \infty$$

for every f in H (see p32, Young[1980]).

Most wavelet books use (3) as the definition of Riesz basis, and call $\{f_n\}$ a Riesz basis with **Riesz bounds** A and B.

Appendix B. Fourier Transform and Fourier Series.

B.1. The L^1 Theory of Fourier Transform.

For $f \in L^1$ the Fourier transform of f is the function $\hat{f}$ defined by

$$\hat{f}(\omega) = (\mathcal{F}f)(\omega) = \int_{-\infty}^{+\infty} f(t)\, e^{-i\omega t} dt$$

for all $\omega \in R$. Some of the basic properties of $\hat{f}(\omega)$ are summarized in the following.

Theorem B1. Let $f \in L^1$. The Fourier transform $\hat{f}$ satisfies

i. The mapping $f \to \hat{f}$ is a bounded linear transform from $L^1(R)$ to $L^\infty(R)$. In fact $\|\hat{f}\|_\infty \leq \|f\|_1$.

ii. If $f \in L^1$ then $\hat{f}$ is uniformly continuous.

iii. $\hat{f}(\omega) \to 0$, as $\omega \to \pm\infty$.

iv. If the derivative f' of f also exists and is in $L^1(R)$, then

$$\hat{f'} = i\omega \hat{f}(\omega)$$

(see p25, Chui [1992a]).

Although $\hat{f}(\omega)\to 0$, as $\omega\to\pm\infty$, it doesn't mean $\hat{f}$ necessarily belongs to $L^1(R)$.

Example B1. The function

$$f(x)=\begin{cases} e^{-x} & x\geq 0 \\ 0 & x<0 \end{cases}$$

is in $L^1(R)$, but its Fourier transform

$$\hat{f}(\omega)=\frac{1}{1-i\omega}$$

is not in $L^1(R)$.

So, generally f can't be "recovered" from $\hat{f}$, but if it happens that $\hat{f}$ is also in $L^1(R)$, then the following holds.

Theorem B2. If f and $\hat{f}$ both belong to $L^1(R)$. Then

$$f(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega t}\hat{f}(\omega)\,d\omega$$

at every point where f is continuous (see p26, Chui [1992a]).

Motivated by the above theorem, the following definition of inverse Fourier transform is formed.

Definition B1. If $f,\hat{f}\in L^1(R)$, then the inverse Fourier transform of $\hat{f}$ is defined by

$$(\mathcal{F}^{-1}\hat{f})(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{i\omega t}\hat{f}(\omega)\,d\omega$$

B.2 The L^2 Theory of Fourier Transform (Plancherel Theory)

The integral defining the Fourier transform is not defined in the Lebesgue sense for all functions in $L^2(R)$. Nevertheless, the Fourier transform has a natural definition on $L^2(R)$ and an elegant theory.

The following theorem is instrumental.

Theorem B3. If $f\in L^1(R)\cap L^2(R)$, then the Fourier transform $\hat{f}$ of f is in $L^2(R)$, and satisfies the following "Parseval Identity"

$$\|\hat{f}\|_2^2=2\pi\|f\|_2^2.$$

(see p16, Stein [1971]).

This theorem asserts that the Fourier transform $\mathcal{F}$ is a bounded linear operator defined on the $L^1(R)\cap L^2(R)$ with range in $L^2(R)$. Since $L^1(R)\cap L^2(R)$ is dense in $L^2(R)$, $\mathcal{F}$ has a bounded extension to all of $L^2(R)$. More precisely, if $f\in L^2(R)$, then its truncation

$$f_N(x)=\begin{cases}f(x) & for\ |x|\le N\\ 0 & otherwise\end{cases}$$

where N=1,2,..., are in $L^1(R) \cap L^2(R)$, so $\hat{f}_N \in L^2(R)$. It is easy to prove $\{\hat{f}_N\}$ is a Cauchy sequence in $L^2(R)$ (regarding convergence in norm or strong convergence), so by the completeness of $L^2(R)$ space, there is a function $\hat{f}_\infty \in L^2(R)$, such that

$$\lim_{N\to\infty} \|\hat{f}_N - \hat{f}_\infty\|_2 = 0$$

Here $\hat{f}_\infty$ is defined as the Fourier transform of $f \in L^2(R)$. The notation $\hat{f} = \mathscr{F} f$ is also used for the Fourier transform of f whenever $f \in L^2(R)$.

Of course, the definition of $\hat{f}$ should be independent of sequences convergent to f. In general, the Fourier transform of $f \in L^2$ is defined as the L^2 limit of the sequence $\{\hat{h}_k\}$ where $\{h_k\}$ is any sequence in $L^1 \cap L^2$ converges to f in the L^2 norm. It can be proved that $\hat{f}$ is independent to the choice of $\{\hat{h}_k\}$.

The Parseval identity can also be extended to $L^2(R)$.

Theorem B4. For all $f, g \in L^2(R)$, the following is true

$$\langle \hat{f}, \hat{g} \rangle = 2\pi \langle f, g \rangle$$

In particular, $\|\hat{f}\|_2 = (2\pi)^{1/2} \|f\|_2$.

Proof: By the definition of $\hat{f}$ in $L^2(R)$ and theorem B3. ▌

Unlike in $L^1(R)$ situation, where the inverse Fourier transform $\mathcal{F}^{-1}$ had to be restricted to the intersection of $L^1(R)$ with the image of $\mathcal{F}$ because $\mathcal{F}$ dos not map $L^1(R)$ to $L^1(R)$. In $L^2(R)$, $\mathcal{F}$ is one-one and onto. So $\mathcal{F}^{-1}$ can easily formulated. In fact, the following theorem holds.

Theorem B5. The Fourier transform $\mathcal{F}$ is a one-one map of $L^2(R)$ onto itself. In other words,, to every $g \in L^2(R)$, there is a unique f such that $\hat{f}=g$ (see p17, Stein [1971]).

By the above theorem, the definition of the inverse Fourier transform of g in L^2 is $f \in L^2$ such that $g=\mathcal{F}f$.

B.3. Fourier Series.

The Fourier series deal with 2π-periodic functions. For each p, $1 \le p < \infty$, let $L^p(0,2\pi)$ denote the class of measurable functions f such that $f(x+2\pi)=f(x)$ and the Lebesgue integral

$$\int_0^{2\pi} |f(x)|^p dx$$

is finite. Also, let $L^\infty(0,2\pi)$ be the class of functions of bounded almost everywhere (a.e.). And $C[0,2\pi]$ denote the 2π-periodic functions which are continuous on $[0,2\pi]$.

Endowed with the norm

$$\|f\|_p=\begin{cases}\int_0^{2\pi}|f(x)|^p dx^{1/p} & for\ 1\le p<\infty\\ \underset{-\infty<x<\infty}{esssup}|f(x)| & for\ p=\infty\end{cases}$$

each $L^p(0,2\pi)$, $1\le p\le\infty$ is a Banach space.

If $f,g\in L^2(0,2\pi)$, using the inner product

$$\langle f,g\rangle=\int_0^{2\pi}f(x)\overline{g(x)}dx$$

$L^2(0,2\pi)$ becomes a Hilbert space.

The Fourier series of a function is defined as

$$\sum_k c_k e^{ikt}$$

where

$$c_k=\frac{1}{2\pi}\int_0^{2\pi}f(t)\,e^{-ikt}dt$$

the partial sum of the Fourier series is denoted by $[S_n(f)](t)$

$$[S_n(f)](t)=\sum_{k=-n}^{n}c_k e^{ikt}$$

Theorem B6. Let $f\in C[0,2\pi]$ and

$$\int_0^a\frac{\sigma(f,t)}{t}dt<\infty$$

for some a>0, where

$$\sigma(f,t)=\sup_{0<h\le t}\max_{x}|f(x+h)-f(x)|,\quad for f\in C[0,2\pi]$$

Then the Fourier series of f converges uniformly to f, i.e.

$$\lim_{n\to\infty}\|f(t)-S_n f\|_{L^\infty[0,2\pi]}=0$$

(p51, vol II, Oryang[1982]).

The above result is called Dini-Lipschitz test of convergence. Before going any further, the notion of bounded variation is introduced.

Let f be a function defined on [a,b], and let $a=x_0<x_1<...<x_n=b$ be any subdivision of [a,b]. Define

$$t=\sum_{i=1}^{n}|f(x_i)-f(x_{i-1})|$$

and

$$T=\sup t$$

T is called the total variation of f over [a,b]. If T<∞, f is said to be of **bounded variation (BV)** over [a,b], and write

$$T=\bigvee_{a\le x\le b} f$$

So, conceptually, a function is of BV means it does not oscillate too drastically.

If one of a or b or both of them are ∞, the total variation and bounded variation over an infinite interval can be defined in the limit.

The following are two important properties regarding functions of bounded variation.

Theorem B7. If f is integrable on [a,b], then the function F defined by

$$F(x)=\int_a^x f(t)\,dt$$

is of continuous function of bounded variation on [a,b].

Proof. To show F is of bounded variation, let $a=x_0<x_1<\ldots<x_k=b$ be any subdivision of [a,b]. Then

$$\sum_{i=1}^{k}|F(x_i)-F(x_{i-1})|=\sum_{i=1}^{k}\left|\int_{x_{i-1}}^{x_i} f(t)\,dt\right|$$

$$\leq\sum_{i=1}^{k}\int_{x_{i-1}}^{x_i}|f(t)|dt=\int_a^b|f(t)|dt.$$

Thus

$$\bigvee_{[a,b]} F \leq \int_a^b |f(t)| dt < \infty \qquad \blacksquare$$

Theorem B8. If f is of bounded variation and g is bounded, then fg is of bounded variation.

Proof. Suppose $x_0 < x_1 < \ldots < x_n$ is any subdivision of the interval, then

$$t = \sum_{i=1}^{n} |f(x_i) g(x_i) - f(x_{i-1}) g(x_{i-1})|$$

$$\leq M \sum_{i=1}^{n} |f(x_i) - f(x_{i-1})|$$

$$\leq M \cdot \bigvee_{x} f$$

where M is a bound of g. $\blacksquare$

The following is famous Dirichlet-Jordan test for convergence.

Theorem B9. If f(t) is of bounded variation on $[-\pi, \pi]$, then the partial sum of its Fourier series

$$\sum_{n=-K}^{+K} c_n e^{i\omega n}$$

is uniformly bounded and converges to $(f(t^+)+f(t^-))/2$ everywhere, that is

$$\lim_{n\to\infty}(S_n f)(t)=\frac{f(t^+)+f(t^-)}{2}$$

for every $t\in R$. Further more, if f is also continuous on any compact interval [a,b], then the Fourier series of f converges uniformly to f on [a,b] (p56, vol II, Oryang[1982]).

While the Fourier transform is defined on $L^p(R)$, the Fourier series only deals with periodic functions. To periodize a function $f\in L^p(R)$, the easiest way is to consider

$$\Phi_f(x)=\sum_{k=-\infty}^{\infty} f(x+2\pi k)\ .$$

The first question arises is whether Φ_f is a function. The answer is positive for p=1, as shown in the following.

Theorem B10. Let $f\in L^1(R)$, then $\Sigma_k f(x+2\pi k)$ converges a.e. to some 2π-periodic function Φ_f. Furthermore, the a.e. convergence is absolute, and $\Phi_f\in L^1(0,2\pi)$.

Proof. Notice the fact

$$\int_0^{2\pi}\Big|\sum_{k=-\infty} f(x+2\pi k)\Big|dx\le\sum_{k=-\infty}\int_0^{2\pi}|f(x+2\pi k)|dx$$

$$\int_0^{-}\Big|\sum_{k=-\infty} f(x+2\pi k)\Big|dx\le\sum_{k=-\infty}\int_0^{--}|f(x+2\pi k)|dx$$

$$\leq \sum_{k=-\infty}^{\infty} \int_{-2\pi k}^{2\pi(k+1)} |f(x)|dx \leq \int_{-\infty}^{\infty} |f(x)|dx < \infty$$

∎

Theorem B11. If $\hat{f} \in L^1(R)$, and f is defined by the inverse Fourier transform, i.e.

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \hat{f}(\omega)\, d\omega$$

then the Fourier series of the 2π-periodic function $\Phi(\omega) = \Sigma_k \hat{f}(\omega + 2k\pi)$ is

$$\sum_{k=-\infty}^{\infty} f(k)\, e^{-i\omega k}$$

Further more, if Φ is of bounded variation on $[0, 2\pi]$, then the following "Poisson Summation Formula" holds:

$$\sum_{k=-\infty}^{\infty} \hat{f}(\omega + 2\pi k) = \sum_{k=-\infty}^{\infty} f(k)\, e^{-i\omega k}$$

at every point where Φ is continuous.

Proof. If the first statement is true, then the next statement is just a simple application of theorem B9.

To establish the first statement, calculate the Fourier coefficient c_k of Φ,

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} e^{-ik\omega}\sum_{j=-\infty}^{\infty}\hat{f}(\omega+2\pi j)\,d\omega$$

$$= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\int_0^{2\pi} e^{-ik\omega}\hat{f}(\omega+2\pi j)\,d\omega$$

$$= \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\int_{2\pi j}^{2\pi(j+1)} e^{-ik\omega}\hat{f}(\omega)\,d\omega$$

$$= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{i\omega(-k)}\hat{f}(\omega)\,d\omega = f(-k)$$

where the interchange of summation and integration is due to the fact that

$$\int_0^{2\pi}\left|\sum_{j=-N}^{M}\hat{f}(\omega+2\pi j)\right|d\omega \le \sum_{j=-N}^{M}\int_0^{2\pi}|\hat{f}(\omega+2\pi j)|\,d\omega$$

$$\le \int_{-2N\pi}^{2\pi(M+1)}|\hat{f}(\omega)|\,d\omega \le \int_{-\infty}^{\infty}|\hat{f}(\omega)|\,d\omega < \infty$$

So Lebesgue Dominated convergence theorem can be applied here.

Hence

$$\sum_{k=-\infty}^{\infty}\hat{f}(\omega+2\pi k) = \sum_{k=-\infty}^{\infty} c_k e^{i\omega k} = \sum_{k=-\infty}^{\infty} f(-k)\,e^{i\omega k} = \sum_{k=-\infty}^{\infty} f(k)\,e^{-i\omega k}$$

at every point where $\Phi(\omega)$ is continuous. ∎

Remark. Most the Fourier analysis books regard the equation

$$\sum_{k=-\infty}^{\infty} f(t+2\pi k) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \hat{f}(k)\, e^{ikt}$$

as the Poisson Summation Formula. The proof of the above theorem is almost identical to the 'usual' Poisson summation formula.

BIBLIOGRAPHY

Carleman, T. [1967] "L'intégrale de Fourier: et questions qui s'y rattachent," leçons professées à l'institute Mittag-Leffler. Uppsala, Paris.

Chui, Charles [1992a]. "An introduction to Wavelets, Vol I in Wavelet analysis and its application." Academic Press, San Diego.

Chui, Charles (Ed) [1992b]. "Wavelets: A Tutorial Theory and Applications, Vol II in Wavelet analysis and its Application". Academic Press, San Diego.

Giardina, Charles R. [1983] Convergence of Discrete Fourier Approximations and Cardinal Expansions. *Technical report 8316, Stevens Institute of Technology.*

Hunt, Richard A. [1976] Developments related to the a.e. convergence of Fourier series. In "Studies in Harmonic Analysis," MAA studies in Mathematics (J. Marshall Ash ed.). The Mathematical Association of America, Rhode Island.

Lathi, B.P. [1989]. "Modern Digital and Analog Communication Systems", 2nd ed. Holt, Rinehart and Winston, Orlando.

Lucky R.W., Salz J. and Weldon, E.J. Jr. [1968]. "Principles of Data Communication." McGraw-Hill, New York.

Körner, T.W [1988]. "Fourier Analysis." Cambridge Univ. Press, London.

Meyer, Yves [1993]. "Wavelets: Algorithms and Applications," translated by Robert Ryan. SIAM, Philadelphia.

Nashed, M. Z. and Walter, G. G. [1991]. General Sampling Theorems for Functions in Reproducing Kernel Hilbert Spaces. *Math. Control Signal Systems*. 4: 363-390.

Oryang, G., Zu, X. and Qing, Z. [1982] "Mathematical analysis," vols I and II. Fudan University Press, Shanghai (in Chinese).

Papoulis, Athanasios [1977]. "Signal Analysis." McGraw-Hill, New York.

Royden, H.L. [1968]. "Real Analysis," 2nd ed. Macmillan, New York.

Rudin, Walter [1964]. "Principle of Mathematical Analysis," 2nd ed. McGraw-Hill, New York.

Rudin, Walter [1991]. "Functional analysis," 2nd ed. McGraw-Hill, New York.

Shapiro, Harold S. [1971]. "Topics in Approximation Theory." Springer-Verlag, Berlin.

Shannon, Claude E. [1949]. Communication in the Presence of Noise. *Proc. I.R.E.* 37, 10-21.

Stein, Elias M. and Weiss, Guido [1971]. "Introduction to Fourier Analysis on Euclidean Spaces." Princeton Univ. Press, Princeton.

Strömberg, Jan-Olov [1991]. A modified Franklin system as the first orthonormal system of wavelets. In "Wavelets and Applications", Y. Meyer, ed. Springer-Verlag, Massou.

Walter, Gilbert [1992]. A Sampling Theorem for Wavelet Subspaces. *IEEE Trans. on Information Theory*. Vol 38, No 2, 881-884.

Zygmund, Antoni [1976]. Notes on the history of Fourier series. In "Studies in Harmonic Analysis," MAA studies in Mathematics (J. Marshall Ash ed.). The Mathematical Association of America, Rhode Island.