5 Wavelet and Gabor Systems

5.1 Invariants of Fourier Analysis and Signal Decompositions

We present wavelet and Gabor systems as a natural development of the basic invariants of Fourier analysis.

a. Define invariants in terms of Theorem 1.2.1
b. Use exercise 1.49 and refer to the metaplectic group again.
c. See Stein in Ash for $\mathbb{R}^d$–invariants.

Definition 5.1.1 (Gabor and Wavelet Systems)

a. Let $g \in L^2(\mathbb{R})$ and let $a, b > 0$.

i. The **Gabor** or **Weyl-Heisenberg system (of coherent states)** is the sequence $\{\phi_{m,n} : (m,n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

(1) $$\phi_{m,n}(t) = e^{2\pi i t m b} g(t - na) = e_{mb}(t)\, \tau_{na} g(t)$$

Clearly,

$$(\phi_{m,n})^\wedge(t) = e^{2\pi i n a m b}\, e^{-2\pi i n a \gamma} \widehat{g}(\gamma - mb) = \tau_{mb}(e_{-na}\widehat{g})(\gamma),$$

cf., *Exercise 1.11* and *Exercise 1.49.*

ii. The **Zak transform** Zg of g is

$$Zg(x,\omega) = a^{1/2} \sum_{k \in \mathbb{Z}} g(xa + ka) e^{2\pi i k\omega}$$

As an exercise, prove that Zg is quasi-periodic (defined in *Example 3.1.10.a*); also try to prove the other properties of Z as given in [BW94].

b. Let $\psi \in L^2(\mathbb{R})$. The **wavelet** or **affine system** corresponding to ψ is the sequence $\{\psi_{m,n} : (m,n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

$$\psi_{m,n}(t) = 2^{m/2} \psi\left(2^m t - n\right).$$

Clearly,

$$(\psi_{m,n})^\wedge(\gamma) = 2^{-m/2} e^{-2\pi i n(\gamma/2^m)}\, \widehat{\psi}(\gamma/2^m) = 2^{-m/2}\, e_{-n} \widehat{\psi}(y/2^m),$$

e.g., *Exercise 5.1.*

Two seminal treatises on wavelet systems, especially wavelet orthonormal bases, are due to Daubechies [Dau92] and Meyer [Mey90].

a. The following are various types of signal decompositions.
b. Recall that Hamel bases were mentioned in Exercise 2.20. Should I define them in 5.5 or 5.1 (or in exercise for ch. 5)?

Definition 5.1.2 (Bases and Frames)

Let H be a Hilbert space and let $\{x_n : n \in \mathbb{Z}\} \subseteq H$ be a sequence in H.

a. The sequence $\{x_n\}$ is a **basis** or **Schauder basis** for H if each $x \in H$ has a unique decomposition

$$x = \sum_{n \in \mathbb{Z}} c_n(x)\, x_n \quad \text{in } H.$$

A basis $\{x_n\}$ for H is an **orthonormal basis (ONB)** for H if it is **orthonormal**, that is, if

$$\forall m, n \in \mathbb{Z}, \quad \langle x_m, x_n \rangle = \delta(m, n).$$

b. A basis $\{x_n\}$ for H is an **unconditional basis** for H if

$$\exists C > 0 \text{ such that } \forall F \subseteq \mathbb{Z}, \text{ where card } F < \infty, \text{ and } \forall b_n, c_n \in \mathbb{C}, \text{ where } n \in F \text{ and } |b_n| < |c_n|,$$

$$\left\| \sum_{n \in F} b_n x_n \right\| \leq C \left\| \sum_{n \in F} c_n x_n \right\|.$$

An unconditional basis is a **bounded unconditional basis** for H if

$$\exists A, B > 0 \quad \text{such that} \quad \forall n \in \mathbb{Z}, \quad A \leq \|x_n\| \leq B.$$

a. Mention noise reduction vav frames vav overcompleteness.
b. Relate frame condition to FS Parseval formula.
c. Use terminology from BHW JFAA on ApproxParseval.

c. The sequence $\{x_n\}$ is a **frame** for H if $\exists A, B > 0$ such that

$$\forall x \in H, \quad A\,\|x\|^2 \leq \sum_{n \in \mathbb{Z}} |\langle x, x_n \rangle|^2 \leq B\,\|x\|^2.$$

The constants A and B are **frame bounds,** and a frame is **tight** if $A = B$. A frame is an **exact frame** if it is no longer a frame whenever any of its elements is removed.

a. Make examples and exercises for above concepts from $\mathbb{R}^d$.
b. Do Schauder bases for Banach spaces.

The **frame operator** of a given frame $\{x_n\}$ for H is the mapping $S : H \to H$ defined by

$$Sx = \sum_{n \in \mathbb{Z}} \langle x, x_n \rangle\, x_n.$$

d. If $\{x_n\}$ is a frame for H and S is the corresponding frame operator, then it can be shown that S is a bijective continuous linear map, and, hence, by the Open Mapping Theorem (*Theorem B9*), S^{-1} is also continuous, e.g. [BW94, Theorem 3.2].

Assuming this result, we can deduce that $\{S^{-1}x_n\}$ is a frame for H with frame bounds B^{-1} and A^{-1}, and that

$$\forall x \in H, \quad x = \sum_{n\in\mathbb{Z}} \langle x, S^{-1}x_n\rangle\, x_n = \sum_{n\in\mathbb{Z}} \langle x, x_n\rangle\, S^{-1}x_n. \tag{2}$$

Using the continuity of S and S^{-1}, the verification of (2) is elementary. In fact,

$$x = SS^{-1}x = \sum_{n\in\mathbb{Z}} \langle S^{-1}x, x_n\rangle\, x_n$$

and

$$x = S^{-1}Sx = S^{-1}\left(\sum_{n\in\mathbb{Z}} \langle x, x_n\rangle\, x_n\right) = \sum_{n\in\mathbb{Z}} \langle x, x_n\rangle\, S^{-1}x_n,$$

Develop that S is positive, tie in with material in HAA Chapter 3, introduce "self-adjoint" (also tie in with HAA Ch. 3?), and thereby completely justify formal calculation in part c. Do *formal* calculations now (for decomp.) and prove in section 5.

a. Check SUP(4) for overlap when assimilate with SUP
b. Use other material from [Ben92b] and [BW94], e.g. from [Ben92b] perhaps include Discussion 15: Vitali's Theorem, Theorem 17: Koethe-Bari-results; from [BW94] perhaps include Definition 3.6 and Theorem 3.7 (to tie in Riesz-basis); from [Ben94, Section 7.4] (not in HAA refs)

e. The notions of bounded unconditional bases and exact frames are equivalent, e.g., [Ben92b, Theorem 17], [BW94, Thereom 3.7], [You80]

Discuss history of frames from Birkhoff, Paley-Wiener, Levinson, Pollard, Riesz-Nagy, Duffin-Schaeffer, DGM (not in HAA refs), Daubechies (some not in HAA). See [Ben92b], [BW94], [Ben94], [Dau92, Chapter 3].

a. Above definition is meaningful for $n \in \mathbb{Z}^d$, and $\mathbb{Z}^d$ is required for $d = 2$ in case of Gabor and wavelet bases and frames.
b. Discuss frame results which will not be proved.

Example 5.1.3 (Signal decompositions: Potpourri and Titillation)

Old History:
Fourier Series (see chapter 3)
Splines (see exercises 1.17 and 2.48)
Bases: Haar, Franklin, Stromberg
Littlewood Paley theory and **Calderon's formula,** [FJW91]

(3) $$f(t) = \int_0^\infty \left(\psi_{1/u} * \psi_{1/u} * f\right)(t)\frac{du}{u}.$$

This formula is a continuous ("overcomplete") wavelet decomposition of $L^2\left(\mathbb{R}^d\right)$. We use the term "wavelet" in the sense of describing f as an integral ("sum") whose harmonics are dilates and translates of a fixed function ψ. In fact, above is

$$\begin{aligned} f(t) &= \int_0^\infty \left(\int \psi_{1/u}(v-t)\,(\psi_{1/u} * f)(v)dv\right)\frac{du}{u} \\ &= \int_0^\infty \int \left(\int \psi_{1/u}(v-w)f(w)dw\right)\psi_{1/u}(v-t)\frac{dv\,du}{u}. \end{aligned}$$

The "sum" is the integral $\int_0^\infty \int \ldots \frac{dv\,du}{u}$ taken over the spatial variable v and the dilations u. The "wavelet coefficients" of f are

$$\int \psi_{1/u}(v-w)f(w)dw, \quad v \in \mathbb{R}^d,\ u > 0.$$

The formal verification of the Calderon's formula is elementary, as long as ψ satisfies certain properties. In fact, the Fourier transform of the right side of (3) is

$$\widehat{f}(\gamma)\int_0^\infty \widehat{\psi}(u\gamma)^2\frac{du}{u};$$

and in *Example 2.2.2* we showed how to construct real-valued, even functions $\psi \in C_c^\infty(\mathbb{R})$ for which

$$\forall\gamma \in (\widehat{\mathbb{R}}) \setminus \{0\}, \quad \int_0^\infty \widehat{\psi}(u\gamma)^2\frac{du}{u} = 1.$$

The extension to $\mathbb{R}^d$ is elementary. Recall that the construction in *Example 2.2.2* also provided finitely vanishing moments of ψ (nota bene: the wavelet intuition that $\widehat{\psi}(0) = 0$ comes from the above requirement; more later).
Quantum mechanics (see [vN55]). Quantum mechanics really deals with Gabor systems in von Neumann's point of view. As such the Heisenberg group arises, just as the $(ax+b)$–group (affine group) underlies wavelet systems (see [Wey5a], as well as the work of Zak and A.J.E.M. Janssen).
Gabor theory (1946) was formulated by Gabor during the same period that he developed holography. The

former was done in terms of the decomposition of acoustical signals, and he received the Nobel prize in physics for the latter. The Heisenberg group plays a fundamental role in both ideas.

A Ph.D. problem/ project would be to quantify my last statement; and to see to what extent Heisenberg group ideas played a role in Gabor's thinking (he didn't know about such groups, but he didn't have to since he was a genius!). Also, the Gabor theory in signal analysis is clearly related to the Heisenberg group.

> **Read Janssen, Wiener Centenary (Symp. in Applied Math 1997), Schempp, and Gabor on such matters.**

Image processing (see Burt and Adelson 1983, as well as A. Rosenfeld in the 1970s)
Speech Processing (late 1970s; see Smith and Barnwell 1986). QMFs and subband coding, with perfect reconstruction, to do compression.
Walsh functions (Walsh 1923). They were used before Walsh by engineers (in an undeveloped form) to reduce channel crosstalk (see NRL-IEEE Symposia on Walsh-functions).
High resolution radar and ambiguity functions
Wigner distributions
Oil (Morlet)

New History
Coifman, Daubechies, Feichtinger, Frazier, Groechenig, Grossmann, Jawerth, Mallat, Yves Meyer, Morlet.

Example 5.1.4 (Gabor's idea)

> **Tie in with UP and refer to HAA pp. 125 and 204; and addend in those places if necessary. Also see 1.1.4. Refer ahead to UP and BLT if I do it in 5.8-5.10**

Gabor's goal was the "best utilization of the information area": the classical Uncertainty Principle inequality is

$$\text{(UP)} \qquad ||f||^2_{L^2(\mathbb{R})} \leq 4\pi\, ||(t-t_0)f(t)||_{L^2(\mathbb{R})}\, ||(\gamma-\gamma_0)\widehat{f}(\gamma)||_{L^2(\widehat{\mathbb{R}})}$$

> **The files, gfo.m and gfl.m, contain information about how to use MATLAB to make pictues like this, and then how to transfer this information to a .tex file.**

Recall "extreme" examples:

a. If $f(t) \sim |t|^a$, $|t| \to \infty$, where $a \in [-3/2, -1/2)$, and $f \in L^2_{loc}(\mathbb{R})$, then $f \in L^2(\mathbb{R})$ (since $a < -1/2$) and

$$||f||^2_{L^2(\mathbb{R})} = \int |t|^2 |f(t)|^2 \, dt = \infty,$$

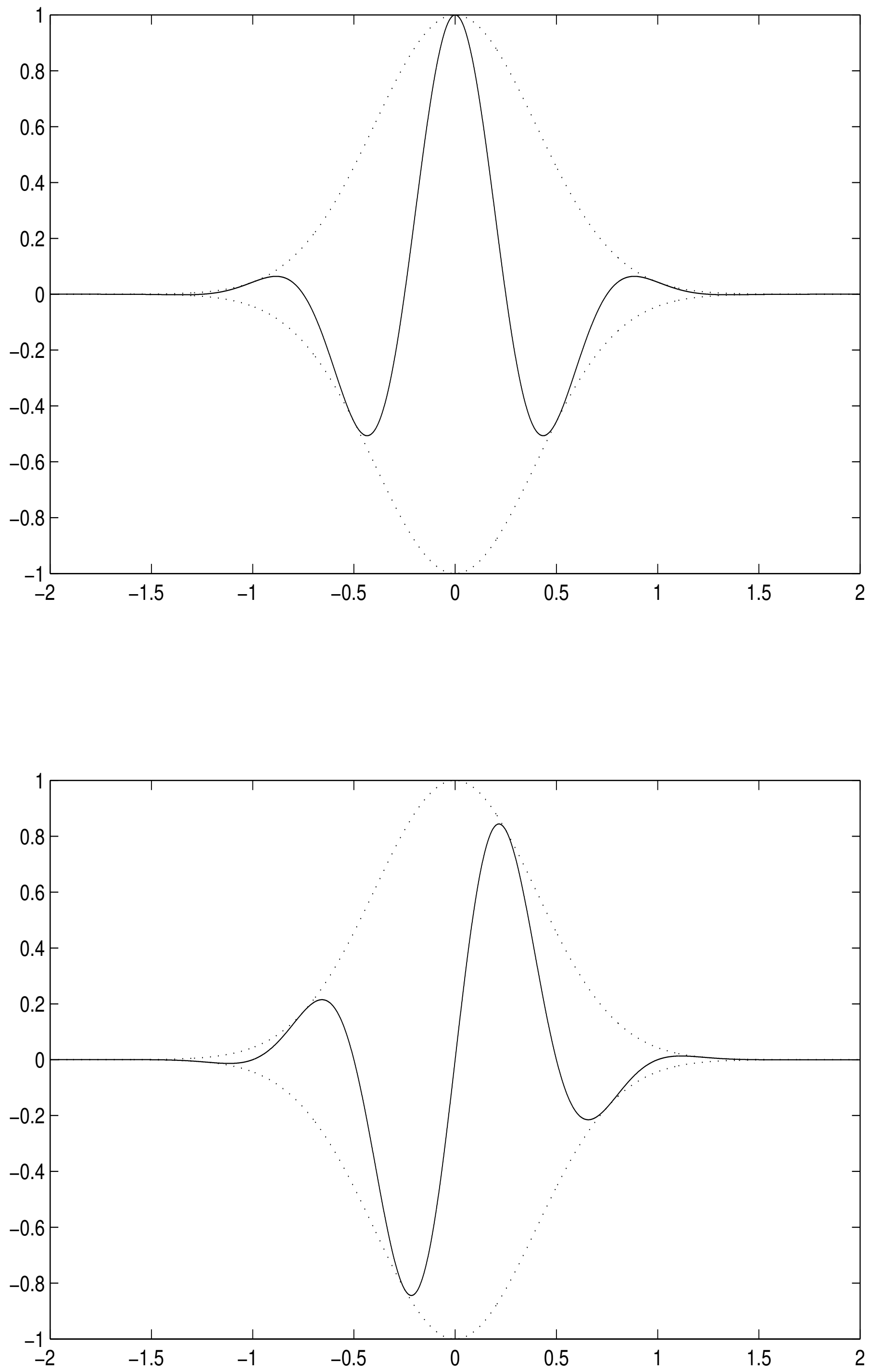


Figure 1: **The real and imaginary parts of a sample (Gaussian) Gabor frame element: $e^{-\pi t^2} cos(2\pi t)$, along with the corresponding Gaussian envelope.**

since
$$\int_A^\infty t^2 t^{2a}\, dt \geq \int_A^\infty t^{2-3}\, dt = \infty$$
(where we have used the hypothesis that $a \geq -3/2$).

b. If $f(t) = \mathbf{1}_{[-T,T)}(\mathrm{t})$, then $\widehat{f}(\gamma) = d_{2\pi T}(\gamma)$, and so $\|\gamma \widehat{f}(\gamma)\|_{L^2(\widehat{\mathbb{R}})} = \infty$.

For $s > 0$ and $(t_0, \gamma_0) \in \mathbb{R} \times (\widehat{\mathbb{R}})$ define
$$g(t) = \sqrt{\frac{2s}{\pi}} e^{-s(t-t_0)^2}\, e^{2\pi i t \gamma_0}.$$
It is an exercise to show that g provides equality in (**UP**).

Define
$$\sigma^2 t_s = 4\pi\, \|(t - t_0) g(t)\|^2_{L^2(\mathbb{R})}$$
and
$$\sigma^2 \gamma_s = 4\pi\, \|(\gamma - \gamma_0)\widehat{g}(\gamma)\|^2_{L^2(\widehat{\mathbb{R}})}\,.$$
As another exercise show that $\sigma^2 t_s = \pi/s$ and $\sigma^2\gamma_s = s/\pi$, so that $\sigma^2 t_s\, \sigma^2 \gamma_s = 1$.

If $f \in L^2(\mathbb{R})$ and $\|f\|_{L^2(\mathbb{R})} = 1$, then the **expected values** associated with $|f|^2$ and $|\widehat{f}|^2$ are
$$\bar{t} = \int t |f(t)|^2\, dt \quad \text{and} \quad \bar{\gamma} = \int \gamma |\widehat{f}|^2\, d\gamma;$$
and the **variances** associated with $|f|^2$ and $|\widehat{f}|^2$ are
$$\sigma^2 t = \int (t - \bar{t})^2 |f(t)|^2\, dt \quad \text{and} \quad \sigma^2 \gamma = \int (\gamma - \bar{\gamma})^2 |\widehat{f}|^2\, d\gamma,$$
resp., see also 2.8.

If $|f|^2$ is a probability density function of a random variable X, then $\sigma^2 t$ is the usual notion of the variance σ_X^2 of X.

The numbers t_0 and γ_0 are not necessarily expected values in the definitions of $\sigma^2 t_s$ and $\sigma^2\gamma_s$; but when they are, these terms present variances.

| **Tie in with 2.8, and addend there if necessary.**

See JB in Wiener Centenary (Proc. Symp. Appl. Math. 1997), as well as [W:M8A, Chapter 7.]

Because of (**UP**) and the equality obtained using above g, Gabor argued that decomposing the time-frequency plane by means of non-overlapping rectangles whose sides each have length and height $\sigma^2 t_s$ and $\sigma^2\gamma_s$, resp., would lead to signal decompositions with optimal time and frequency localization.

Gabor's decomposition is of the form

$$f(t) = \sum_{m,n \in \mathbb{Z}} c_{m,n} \sqrt{\frac{2s}{\pi}} e^{-s(t-n\sigma^2 t_s)^2} e^{2\pi i m \sigma^2 \gamma_s}.$$

Generalizations include wavelets, STFT and frame decompositions, non-uniformly rectilinear dissections of the time-frequency plane and other uncertainty principles.

It is natural to investigate orthogonality for the double sequence $\{g(t-na)e_{mb}(t)\}$.

Perhaps ask a problem on above before giving more results. Also, as an exercise, evaluate the products to measure the degree of orthogonality, see exercise 22, fall 1997

Example 5.1.5 (Morlet's idea)

a. Jean Morlet is a geophysicist who introduced wavelet theory in the course of his analysis of seismic traces s, while employed by Elf Aquitaine.

Reverify names!

References:

i. J. Morlet, Sampling Theory and Wave Propagation, Soc. Exploration Geophysicists, 1981

ii. P. Goupillaud, A. Grossmann and J. Morlet, Cycle Octave Analysis, Geoexploration, 23 (1984/1985), 85-102.

Morlet's idea was to analyze the trace s by a family of functions having a fixed shape, in the sense that each element of the family should be a (lattice) translation and/or (dyadic) dilation of one function ψ. Thus, the elements $\psi_{m,n}$ of the family are indexed by pairs (m,n) of integers, and each $\psi_{m,n}$ is defined as

$$\psi_{m,n}(t) = 2^{m/2}\,\psi\left(2^m t - n\right).$$

The aforementioned analysis is the computation of the sequence $\{\langle s, \psi_{m,n}\rangle\}$, where s and ψ are assumed to be elements of $L^2(\mathbb{R})$, and $\langle s, \psi_{m,n}\rangle \equiv \int s(t)\overline{\psi_{m,n}}\,dt$. The goal is to reconstruct the trace s from the computable data set $\{\langle s, \psi_{m,n}\rangle\}$.

The point of dealing with $\{\psi_{m,n}\}$ is that the operation of dilation and/or translation preserves the *same* number of cycles for high, medium or low frequencies, e.g., see below.

The sampled values $\langle s, \psi_{m,n}\rangle$ are analogous to the Fourier coefficients of s, in that we want to reconstruct s in terms of them. There are, however, fundamental differences between the wavelet system $\{\psi_{m,n}\}$ and Fourier analysis.

Before discussing these differences, and even before explicitly dealing with Morlet's original function ψ, we want to make a remark about terminology:

The term **wavelet** often means a function ψ for which the family $\{\psi_{m,n}\}$ defined by (**??**) is an ONB for $L^2(\mathbb{R})$ (we shall define such functions as **orthonormal wavelets** in the next section.) On the other hand, Morlet's original work, which inspired so much of wavelet theory in the 1980s, was well–understood by Morlet and other geophysicists at the time to be a redundant wavelet system.

As reinforced to us by Goupillaud [1997 AMS Special Session in Wavelet Theory, San Diego], Morlet's family $\{\psi_{m,n}\}$ could not be orthonormal or independent in order to achieve the noise reduction required to solve the geophysical problem at hand.

b. Morlet considered the function

$$\psi(t) = e^{-\pi t^2}\left(e^{2\pi i t\gamma_0} - e^{-\pi\gamma_0^2}\right),$$

e.g., [Dau92, pages 75-78].

For example, the real and imaginary parts of the Morlet wavelet (with $m = n = 0$), are as follows:

> **Perhaps this goal was desired to get rid of DC-component for proc. signals, i.e. to get rid of a huge area**

As an exercise, show that $\widehat{\psi}(0) = 0$.

> **Solution:**
>
> $$\begin{aligned} \widehat{\psi}(0) &= \int e^{-\pi t^2} e^{2\pi i t\gamma_0}\,dt - e^{-\pi\gamma_0^2}\int e^{-\pi t^2}\,dt \\ &= g_{\sqrt{\pi}}(\gamma_0) - e^{-\pi\gamma_0^2}\int g_{\sqrt{\pi}}(t)\,dt \\ &= e^{-\pi\gamma_0^2} - e^{-\pi\gamma_0^2} = 0 \end{aligned}$$

The geometric properties attributed to $\{\psi_{m,n}\}$ in part a. are illustrated by plotting $\operatorname{Re}\psi$ and $\operatorname{Im}\psi$ for various values of γ_0. Note that if $n = 0$ and $\gamma_0 = 1$, then

$$\psi_{m,0}(t) = 2^{m/2}\,e^{-\pi 2^{2m} t^2}\left(e^{2\pi i 2^m t} - e^{-\pi}\right).$$

Plot for values of m both large and small.

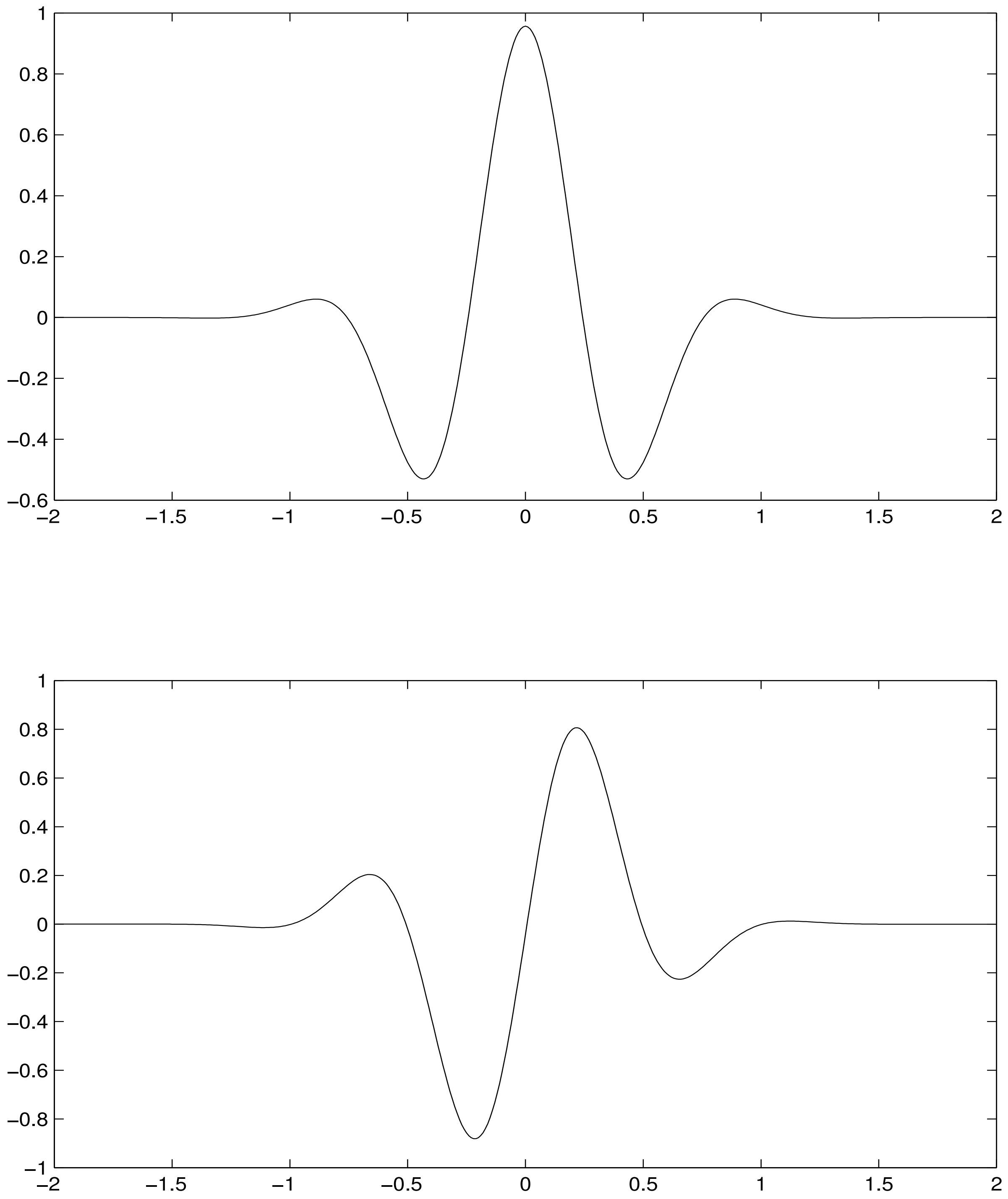


Figure 2: **The Real and Imaginary Parts of the Morlet Wavelet with $m = n = 0$.**

$\widehat{\psi}(0) = 0$, **i.e.,** ψ **has no DC-component, the average of** ψ **is 0.**
(Generally subtract the DC-component to analyze the data in a finer way).

i. **Tell the value of this property in signal processing (DSP) since the integral of positive data could be so large as to distort the activity of the data or make it difficult to read.**

ii. **In wavelet theory, it is desirable to have** $\widehat{\psi}(0) = 0$ **since reconstruction of continuous wavelet transform requires a condition which is satisfied when** $\widehat{\psi}(0) = 0$ **(in a stronger way, e.g.,** $\widehat{\psi} = 0$ **on a n'hood of 0, i.e., the admissibility condition for representation, cf., the calculation following the Calderon formula.)**

iii. **Discuss the frame properties of the Morlet system and put it in context of affine frame criteria of Daubechies [Dau92].**

iv. **If** $\psi, \widehat{\psi} \in L^1 \cap L^2$ **and** $\{\psi_{m,n}\}$ **is ON, then** $\widehat{\psi}(0) = 0$

Exercise (Morlet) continued (also see 5.1.6):

$$Re\psi_{m,0}(t) = 2^{m/2} e^{-\pi 2^{2m} t^2} (cos(2\pi 2^m t\gamma_0) - e^{-\pi\gamma_0^2}).$$

Suppose $\gamma_0 > 0$, and m is large. Clearly, $e^{-\pi t^2} < 1$, $(\pi\gamma_0^2 > 1 \Rightarrow -\pi\gamma_0^2 < 0 \Rightarrow e^{-\pi\gamma_0^2} < 1)$. On the following two pages, the dotted lines above the origin represent the values $e^{\pi\gamma_0^2}$ and $1 - e^{\pi\gamma_0^2}$, on the left and the right columns, respectively.

For seismic traces s, **one considers the function** $f = s + iIm(f) \in H^2$, **where** $f \in H^2$ **means that** $f \in L^2(\mathbb{R})$ **and supp**$(\hat{f}) \subset [0, \infty)$. **See the Grossman-Morlet paper, page 723.**

For the Morlet function in (5.1.5.b), we compute

$$\hat{\psi}(\gamma) = e^{-\pi(\gamma-\gamma_0)^2} - e^{-\pi\gamma_0^2} e^{-\pi\gamma_0}$$

and so $\hat{\psi}_{m,n}(\gamma) =$ (See exercise 23, Fall 1997).

$$e^{-\pi r t^2} \leftrightarrow g_{\sqrt{\pi/r}}\ ,\ g(t) = \frac{1}{\pi} e^{-\pi t^2},$$

$$g_\lambda(t) = \frac{\lambda}{\sqrt{\pi}} e^{-(\lambda t)^2}\ ,\ \int g_\lambda(t)dt = 1\ ,$$

$$g_\lambda(t) \leftrightarrow e^{-(\pi\gamma/\lambda)^2}\ ,$$

$$e^{-\pi t^2} e^{2\pi i t\gamma_0} \leftrightarrow e^{-\pi(\gamma-\gamma_0)^2}$$

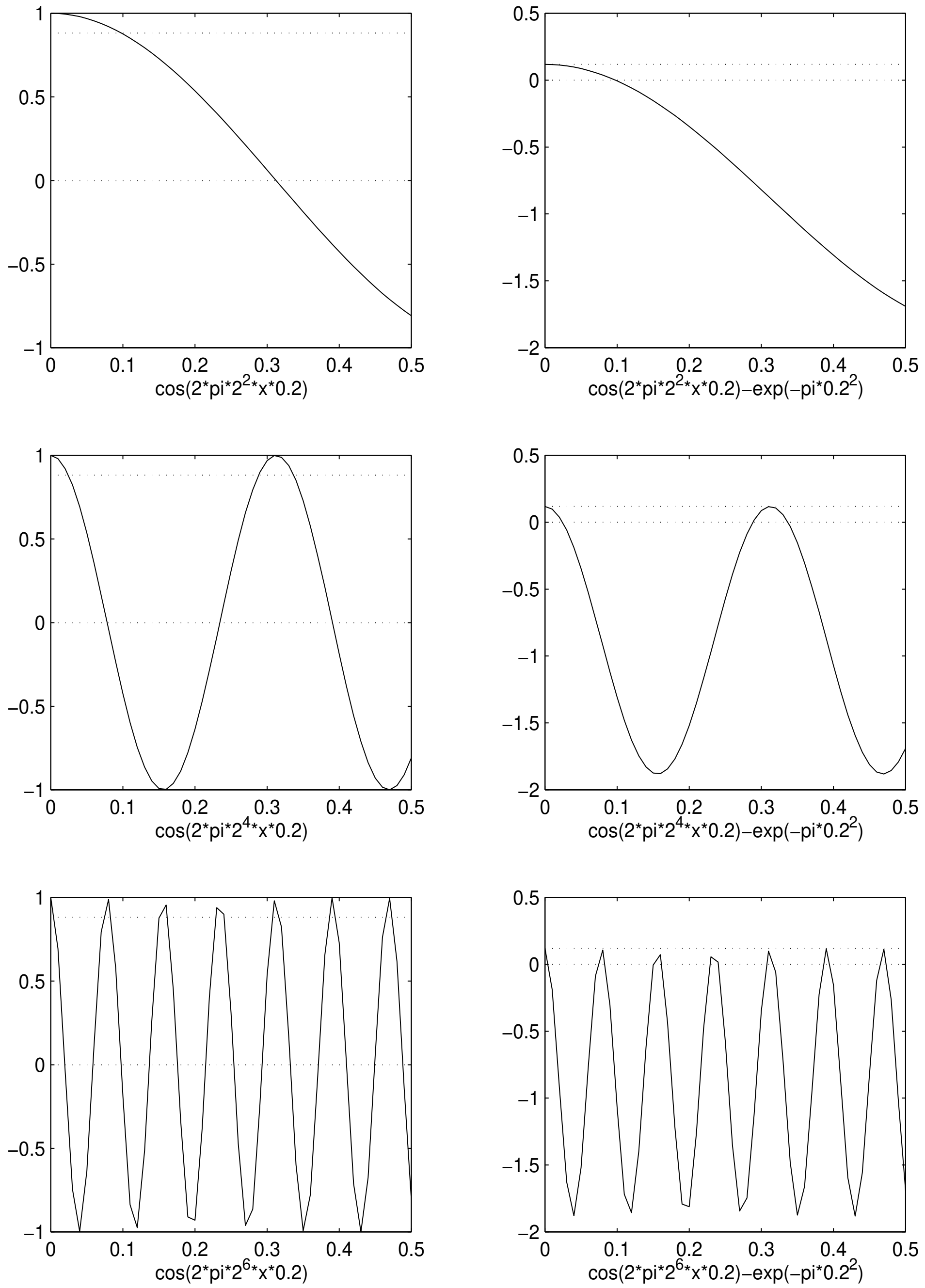


Figure 3: **Examples of the Morlet Cosine Part with $\gamma = .2$**

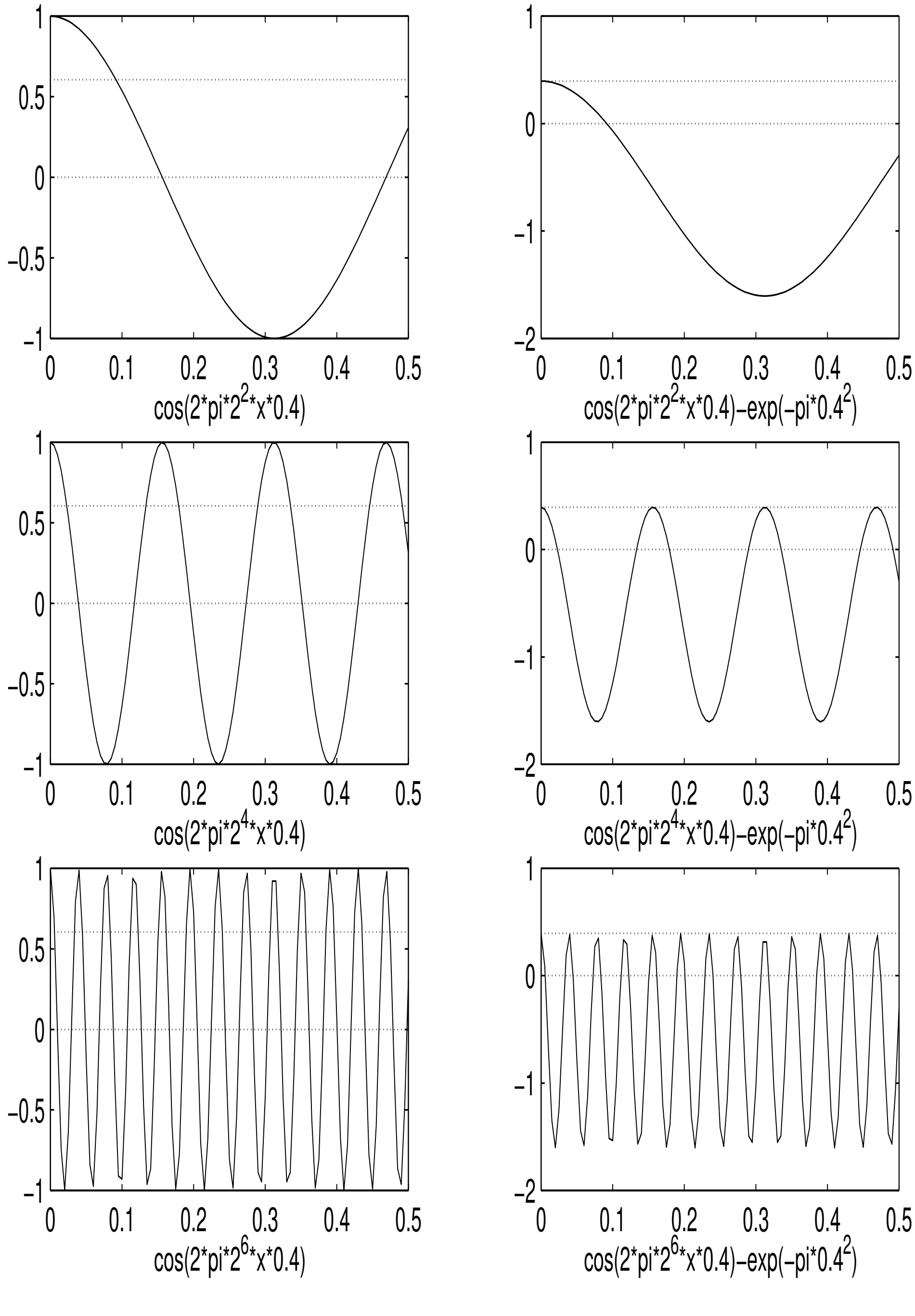


Figure 4: **Examples of the Morlet Cosine, with $e^{-\pi\gamma^2}$ subtracted for $\gamma = .4$**

"In practice, γ_0 is chosen so that the difference between the first and second maxima of ψ is large enough."

Make an *Exercise* based on this remark.

Exercise:

a. Show that $\{\psi_{m,n}\}$ is never orthogonal. (Exercise 13, Fall 1997) b. Are there any subsequences for which there is orthogonality? (Exercise 24, Fall 1997)

We'll see later that if θ, $\hat{\theta} \in L^1 \cap L^2(\mathbb{R})$ and $\{\theta_{m,n}\}$ is orthogonal, then $\int \theta(t)\,dt = 0$.

Stress the types of decompositions sought in the AGU abstracts or perhaps in the Geophysics/wavelets book from that conference.

In [GGm 84, page 90], (Goupillaud, Grossman, Morlet), the "wavelet"

$$\psi_b(t) = e^{2\pi ibt}e^{-\pi t^2/2} - \sqrt{2}e^{-\pi b^2/4}e^{2\pi ibt}e^{-\pi t^2} \tag{1}$$

is used. They define the Fourier transform as $\frac{1}{\sqrt{2\pi}}\int f(t)e^{-2\pi it\gamma}\,dt$. **As such,**

$$\hat{\psi_b}(\gamma) = e^{-(\gamma-b)^2/2} - e^{-\pi b^2/4}e^{-(\gamma-b)^2/4}. \tag{2}$$

[When b is a large given number, ψ is essentially $e^{2\pi ibt}e^{-\pi t^2/2}$ which was the earliest "wavelet" from the 1980s,] cf., [CGT89, pages 10-12].
In [[Dau92, pages 76-78]], Daubechies uses

$$\psi(t) = \pi^{-1/4}e^{-\pi t^2/2}(e^{2\pi ibt} - e^{-\pi b^2/2})^* \tag{3}$$

as the type of "wavelet" associated with the early work of Morlet and Kronland-Martinet, Comber et al, cf., [CGT89], which contains the article *Reading and understanding continuous wavelet transforms* (Grossman, K-M, Morlet) in [CGT89, pages 2-20].
*** In the case of (??),**

$$\psi(\gamma) = \pi^{-1/4}(e^{-(\gamma-b)^2/2} - e^{-\pi b^2/2}e^{-\gamma^2/2}),$$

where Daubechies defines the Fourier transform as above.

Example 5.1.6 *Plots of Gabor and Morlet systems*

Exercise. Make MATLAB figures of Gabor case, both for the Gaussian and perhaps for some other $L^2(\mathbb{R})$ functions.

Draw both the real and imaginary parts, and see if I can integrate them in a visually meaningful way.

See my "Morlet folder" as well as my MATLAB diary called "Morlet". There are 4 MATLAB figures of Morlet functions in these notes as of August 1997.

Exercise: Devise more MATLAB exercises with Morlet. High/low frequences v.a.v. invariance of cycles.

In this section, we present several examples of the Gabor and Morlet functions. Note: For Gabor functions, shifts in m change the number of cycles inside the Gaussian envelope, but do not affect the size of that envelope. Shifts in n, in the Gabor case lead to translations of the original function, but these translations are not dependent on m. For the Morlet functions, we note that changes in m lead to dilations or contractions of the function, but do not change the number of cycles in the Gaussian envelope. The translations caused by changes in n are also different, as they now depend on m, as well. For the Morlet case, $\psi_{m,n}$ is now a diation of the translate of ψ by $n/2^m$. This effect can be seen in the two Morlet examples pages. and noting the shifts of the Morlet functions are smaller at the bottom (where m is larger), and are more pronounced at the top (where m is smaller), even when n is the same in any given column.

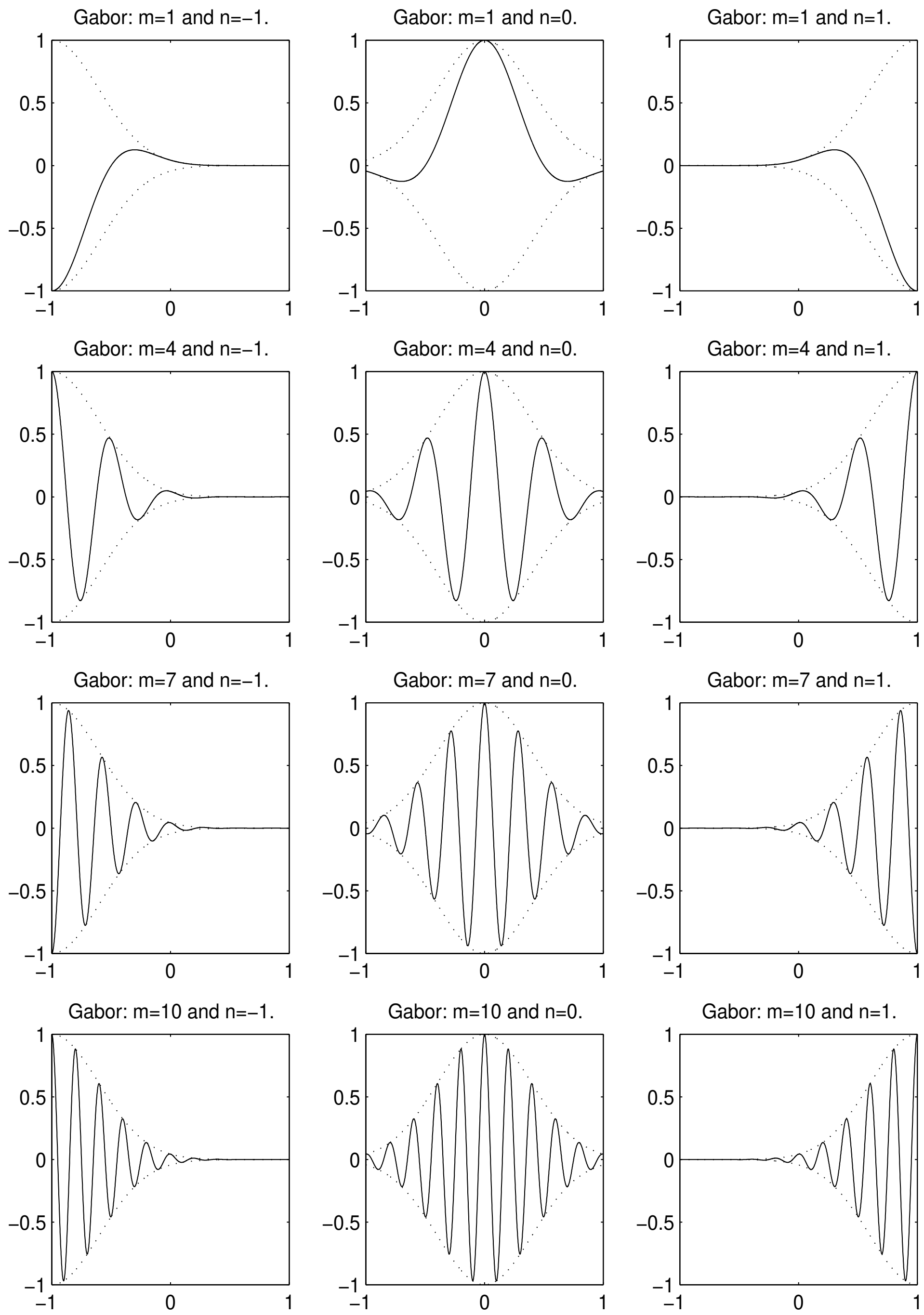


Figure 5: **Gaussian Gabor Frame Elements**

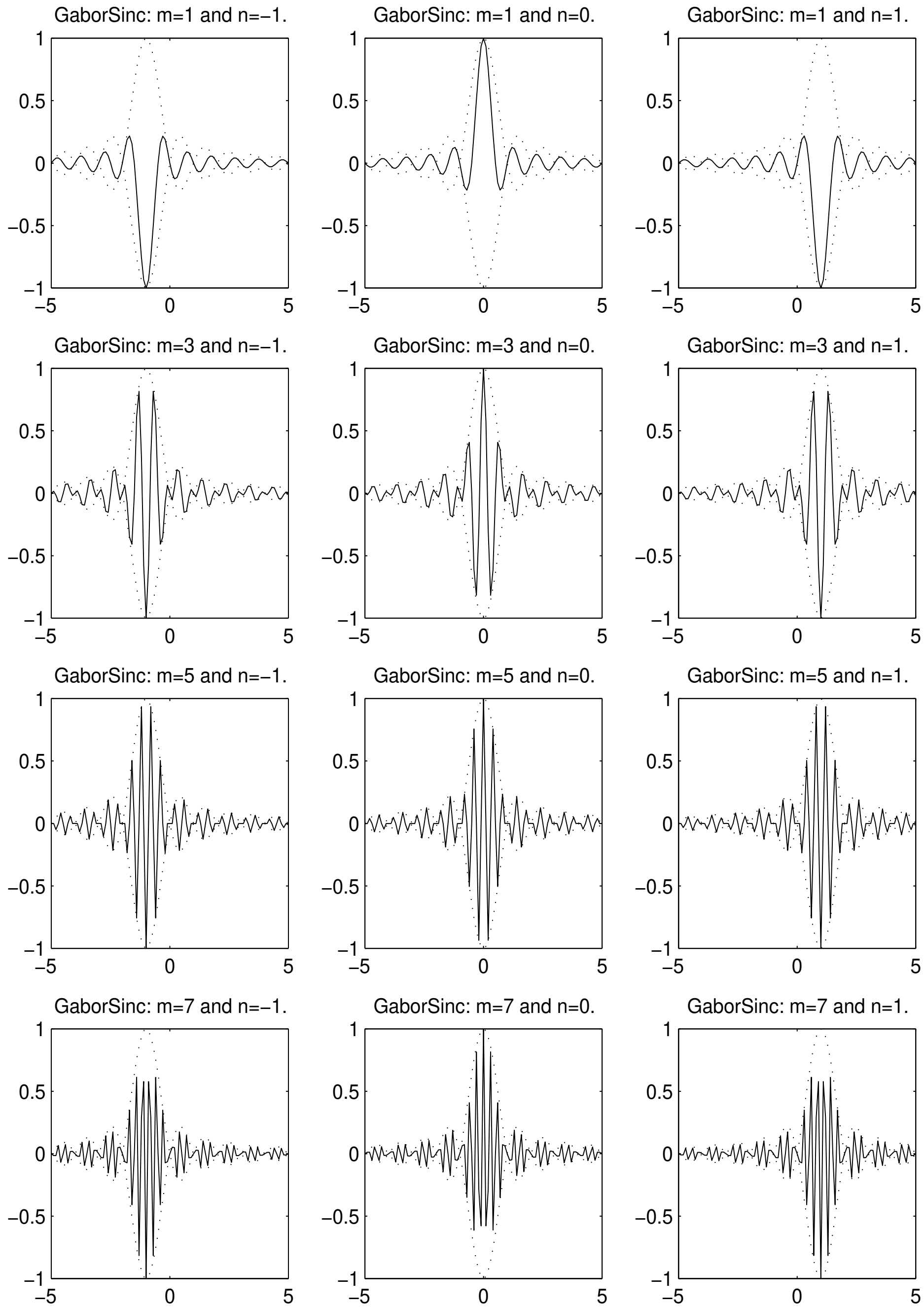


Figure 6: **Sinc Gabor Frame Elements** $(g = sin(\pi t)/(\pi t))$

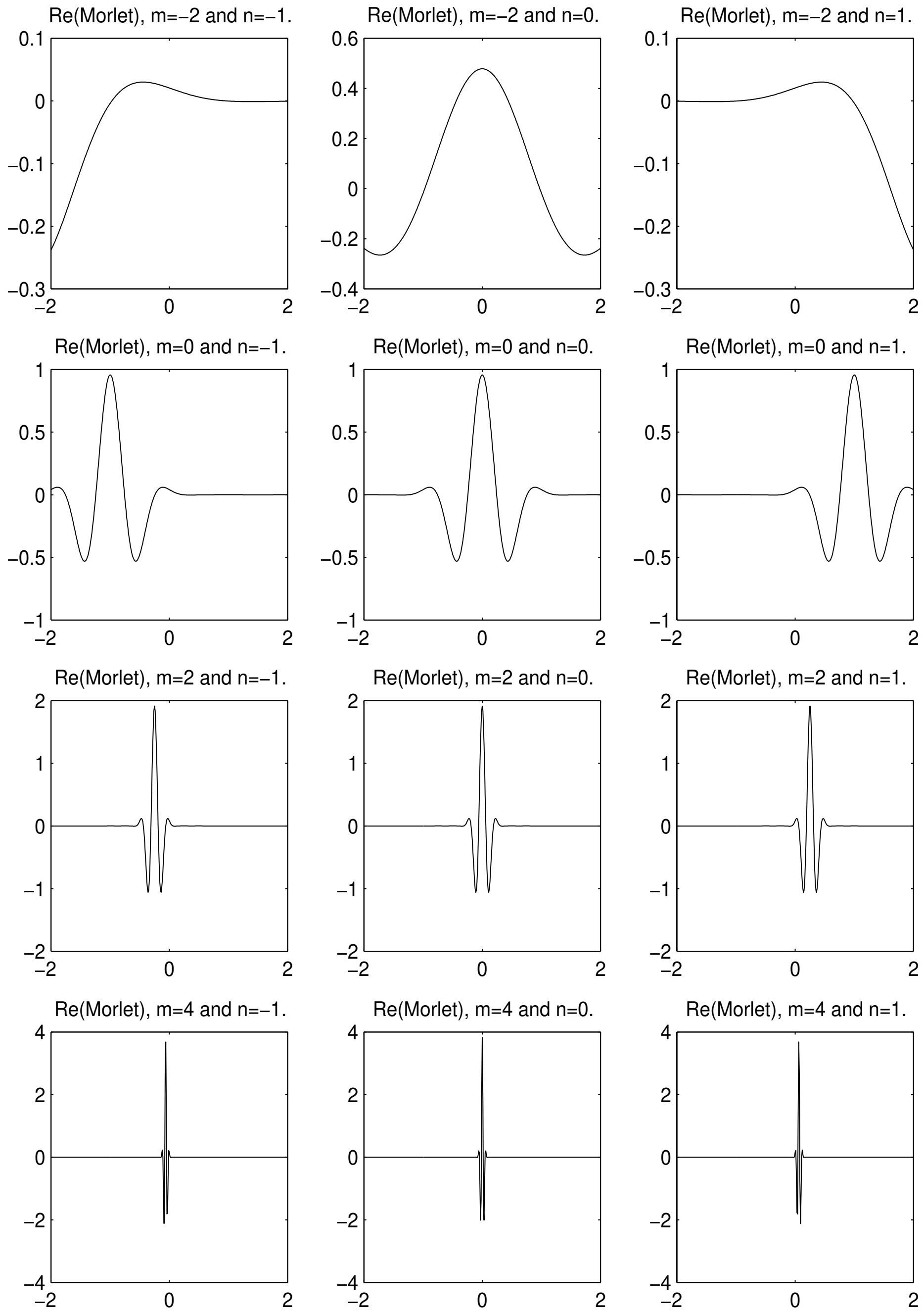


Figure 7: **Real Part of the Morlet Wavelet**

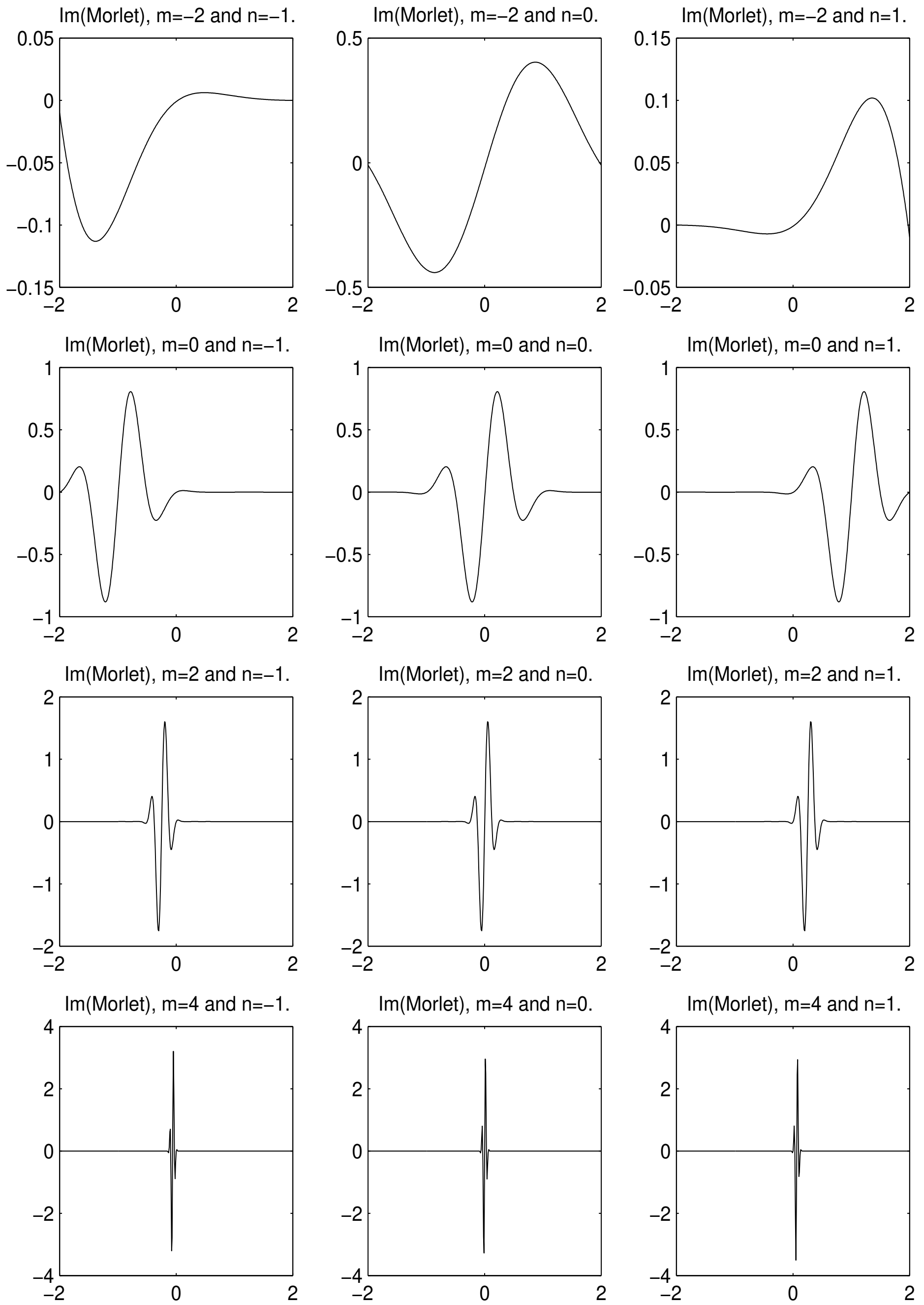


Figure 8: **Imaginary part of the Morlet Wavelet**

Example 5.1.7 (The Shannon wavelet system)

a. Let $\Omega > 0$. To be compatible with standard wavelet notation, let

$$\phi = \phi_{(\Omega)} = \frac{1}{\sqrt{2\Omega}} d_{2\pi\Omega}$$

and let

$$\psi = \psi_{(\Omega)} = \frac{1}{\sqrt{2\Omega}} (d_{2\pi(2\Omega)} - d_{2\pi(\Omega)}),$$

(The parentheses are put around Ω in the definitions of $\phi_{(\Omega)}$ and $\psi_{(\Omega)}$ to distinguish these functions from Ω-dilations.) $\psi_{(\Omega)}$ is the **Shannon wavelet**. [We shall suppress the subscript (Ω) unless confusion arises.]

b. Note that

$$\begin{aligned} \sqrt{2\Omega}\widehat{\psi} &= \mathbf{1}_{[-2\Omega,2\Omega)} - \mathbf{1}_{[-\Omega,\Omega)} \\ (5.1.\text{g.a}) \qquad & \\ &= \mathbf{1}_{[-2\Omega,-\Omega)} + \mathbf{1}_{[\Omega,2\Omega)}, \end{aligned}$$

and hence

$$(5.1.\text{g.a}) \qquad \sqrt{2\Omega}\widehat{\psi}(\gamma) = \mathbf{1}_{[-\Omega,\Omega)}(2\gamma - 3\Omega) + \mathbf{1}_{[-\Omega,\Omega)}(2\gamma + 3\Omega)$$

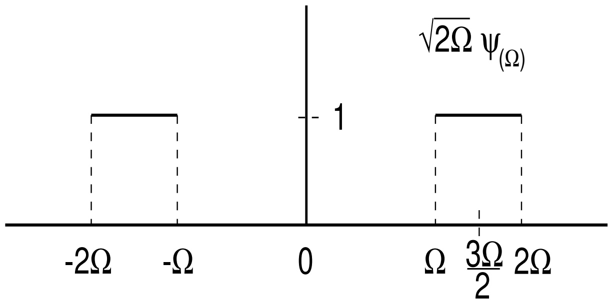
The fact that $\widehat{\psi}(0) = \int \psi(t)dt = 0$ is relevant in wavelet theory.

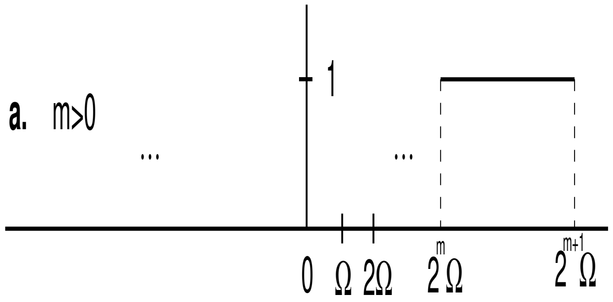
| **See Notes in Section 5.1.e. about $\widehat{\psi} = 0$.**

$\mathbf{1}_{[-\Omega,\Omega)}(2\gamma - 3\Omega) = 1$ if and only if $2\gamma - 3\Omega \in [-\Omega, \Omega)$ if and only if $2\gamma \in [2\Omega, 4\Omega)$, i.e., $\gamma \in [\Omega, 2\Omega)$. Similarly, $\mathbf{1}_{[-\Omega,\Omega)}(2\gamma + 3\Omega) = 1$ if and only if $2\gamma + 3\Omega \in [-\Omega, \Omega)$, i.e., $2\gamma \in [-4\Omega, -2\Omega)$, i.e., $\gamma \in [-2\Omega, -\Omega)$.

| **See** *Exercises 5.1, 5.2* **(Exercise is about Shannon)**

$$\begin{aligned} \psi_{m,n}(t) &= \frac{2^{m/2}}{\sqrt{2\Omega}} \left(d_{2\pi(2\Omega)}(2^m t - n) - d_{2\pi\Omega}(2^m t - n)\right) \\ & \frac{2^{m/2}}{\sqrt{2\Omega}} \frac{\sin 2\pi(2\Omega)(2^m t - n) - \sin 2\pi\Omega(2^m t - n)}{\pi(2^m t - n)}. \end{aligned}$$





Also, by (5.1.g.a) we have

$$\sqrt{2\Omega}\widehat{\psi}(\gamma) = \tau_{\frac{3\Omega}{2}}\mathbf{1}_{[-\frac{\Omega}{2},\frac{\Omega}{2})}(\gamma) + \tau_{-\frac{3\Omega}{2}}\mathbf{1}_{[-\frac{\Omega}{2},\frac{\Omega}{2})}(\gamma)$$

and, thus,

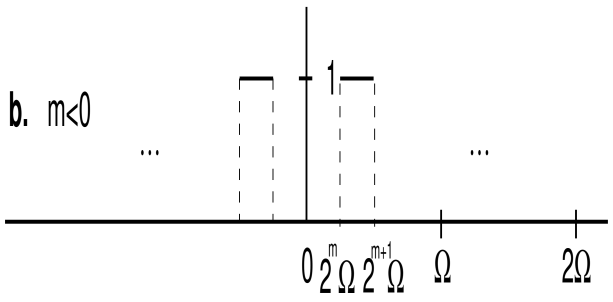
$$\sqrt{2\Omega}\psi(t) = 2d_{2\pi(\frac{\Omega}{2})}(t)cos2\pi(\frac{3\Omega}{2})t.$$

Note that $2d_{2\pi(\frac{\Omega}{2})}(t) = d_{2\pi\Omega}(\frac{t}{2})$.

c. Clearly, $\psi_{0,0} = \psi$. If $m > 0$, resp., $m < 0$, then $\sqrt{2\Omega}\widehat{\psi}(\gamma/2^m)$ is depicted in **Figure a**, resp., **Figure b**.

$\gamma/2^m \in [\Omega, 2\Omega)$ **implies** $\gamma \in [2^m\Omega, 2^{m+1}\Omega)$**, and the length of this latter interval is** $2^m(2\Omega - \Omega) = 2^m\Omega$**.**

$m < 0$ **implies** $m+1 \le 0$ **implies** $m < m+1 \le 0$ **implies** $-m > -(m+1) \ge 0$ **implies** $2^{-m} > 2^{-(m+1)} \ge 1$ **implies** $2^m < 2^{m+1} \le 1$.



Coifman, Meyer, and Wickehauser (Ruskai, et al. page 454) use the term (with quotations) "Shannon wavelet" for the function we have so defined (in the $\Omega = \frac{1}{2}$ case).
In ID's 10 lectures, page 115, ID refers to $\{(\psi_{(1/2)})_{m,n}\}$ as the Littlewood - Paley basis.
It is also called the *sinc wavelet* (in Vetterli - Kovacevic).

$$\begin{aligned}\|\psi_{(\Omega)}\|_{L^2(\mathbb{R})} &= \left\|\widehat{\psi}_{(\Omega)}\right\|_{L^2(\widehat{\mathbb{R}})} \\ &= \left(\frac{1}{2\Omega}\int_{-2\Omega}^{-\Omega} d\gamma + \frac{1}{2\Omega}\int_{\Omega}^{2\Omega} d\gamma\right)^{\frac{1}{2}} \\ &= 1.\end{aligned}$$

d. Note that $\|\phi\|_{L^2(\mathbb{R})} = \|\psi\|_{L^2(\mathbb{R})} = 1$, a fact that in dear by the P-P theorem.

Inset figures of Shannon wavelets, using MATLAB, using $\Omega < 1$ and $\Omega > 1$

As a introduction to Theorem 5.1.h, we want to make a remark to the following effect:

Because of the expectation of reconstructing a bandlimited function as we have done in *Theorem 3.10.10* (Classical Sampling Theorem) and with the point of view of *Definition 3.10.11*, we shall now consider an arbitrary signal $f \in L^2(\mathbb{R})$, and shall prove a reconstruction formula for f by splitting the frequency response $F = \widehat{f}$ into a sequence of functions of bounded bandwidth. The theme is the interplay of Fourier series and Fourier transforms, and the technique invokes the Shannon wavelet.

Sections 3.8 - 3.10 will be part of HAA4.

Draw figures using MATLAB distinguishing $\frac{n}{4\Omega}$ & $\frac{n}{2\Omega}$. Emphasize nonON for $\frac{n}{4\Omega}$.

Theorem 5.1.8 (A Shannon Wavelet Decomposition using Fourier Series)

Let $f \in L^2(\mathbb{R})$ have Fourier transform F, and let $\Omega > 0$. For the function $\phi \equiv \phi_{(\Omega)} \equiv \frac{1}{\sqrt{2\Omega}} d_{2\pi\Omega}$ and the Shannon wavelet $\psi \equiv \psi_{(\Omega)} \equiv \frac{1}{\sqrt{2\Omega}}(d_{2\pi(2\Omega)} - d_{2\pi\Omega})$, there is the decomposition

$$
\begin{aligned}
f &= \sqrt{2\Omega} f * \phi + \sum_{m=0}^{\infty} \sum_n d_{m,n} \psi_{m,n/(4\Omega)} \\
(5.5.\text{h a}) \qquad & \\
&= \sum_n f * \phi(\frac{n}{2\Omega}) \tau_{n/(2\Omega)} \phi + \sum_{m=0}^{\infty} \sum_n d_{m,n} \psi_{m,n/(4\Omega)} \\
&= \sum_m \sum_n d_{m,n} \psi_{m,n/(4\Omega)} \qquad \text{in } L^2(\mathbb{R}) \ ,
\end{aligned}
$$

where

$$
d_{m,n} = \frac{1}{\sqrt{2\Omega} 2^{(m/2)+1}} \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} F(\gamma) \left(\mathbf{1}_{[-2^{m+1}\Omega, -2^m\Omega)}(\gamma) + \mathbf{1}_{[2^m\Omega, 2^{m+1}\Omega)}(\gamma) \right) * e^{2\pi i n\gamma/(2^{m+2}\Omega)} d\gamma
$$

$$
f * \phi(\frac{n}{2\Omega}) = \int f(t) \overline{\phi(\frac{n}{2\Omega} - t)} \ dt = \frac{1}{\sqrt{2\Omega}} \int_{-\Omega}^{\Omega} \hat{f}(\gamma) e^{2\pi i (\frac{n}{2\Omega})\gamma} d\gamma.
$$

$$
\Longrightarrow \quad \sqrt{2\Omega} f * \phi(\frac{n}{2\Omega}) = f(\frac{n}{2\Omega}) \qquad \textbf{if } f \in PW_\Omega
$$

Proof

a. We can write $F \in L^2(\widehat{\mathbb{R}})$ as

$$
\begin{aligned}
(5.1.hc) \qquad F(\gamma) &= \sqrt{2\Omega} F(\gamma) \hat{\phi}(\gamma) + \sqrt{2\Omega} \sum_{m=0}^{\infty} F(\gamma) \hat{\psi}(\frac{\gamma}{2^m}) \\
&= \sqrt{2\Omega} \sum_{m=-\infty}^{-1} F(\gamma) \hat{\psi}(\frac{\gamma}{2^m}) + \sqrt{2\Omega} \sum_{m=0}^{\infty} F(\gamma) \hat{\psi}(\frac{\gamma}{2^m}),
\end{aligned}
$$

where the convergence is in the L^2-norm.

Setting $F_m(\gamma) = \sqrt{2\Omega} F(\gamma) \hat{\psi}(\gamma/2^m)$, we denote the inverse Fourier transform $F_m{}^\vee$ by $f_m \in L^2(\mathbb{R})$. Thus,

$$
\forall m \in \mathbb{Z}, \qquad f_m \in PW_{2^{m+1}\Omega},
$$

and , in particular,

$$
supp \ \ F_m \subseteq [-2^{m+1}\Omega, -2^m\Omega] \cup [2^m\Omega, 2^{m+1}\Omega].
$$

b. We now introduce Fourier series in the following way. Consider F_m as a $2^{m+2}\Omega$-periodic function on $\widehat{\mathbb{R}}$, which is equal to $\sqrt{2\Omega}F(\gamma)\widehat{\psi}(\gamma/2^m)$ on $[-2^{m+1}\Omega, 2^{m+1}\Omega)$. The Fourier series of F_m is

$$S(F_m)(\gamma) = \sum_n c_{m,n} e^{-2\pi i n\gamma/(2^{m+2}\Omega)},$$

where, for each fixed m, $\{c_{m,n} : n \in \mathbb{Z}\} \in l^2(\mathbb{Z})$ and, in fact,

$$\forall n \in \mathbb{Z}, \quad c_{m,n} = \frac{1}{2^{m+2}\Omega}\int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} F_m(\gamma) e^{2\pi i n\gamma/(2^{m+2}\Omega)} d\gamma.$$

By the L^2 theory of Fourier transforms (Section 1.10), we have

$$f_m(t) = \int F_m(\gamma) e^{2\pi i t\gamma} d\gamma \qquad \text{in } L^2(\mathbb{R}),$$

and so

$$\text{(5.1.h d)} \qquad \begin{aligned} f_m(t) &= \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} \sqrt{2\Omega}F_m(\gamma)\widehat{\psi}(\frac{\gamma}{2^m}) e^{2\pi i t\gamma} d\gamma \\ &= \sqrt{2\Omega}\sum_n c_{m,n} \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} \widehat{\psi}(\frac{\gamma}{2^m}) e^{2\pi i \left(t - n/(2^{m+2}\Omega)\right)\gamma} d\gamma \quad \text{in } L^2(\mathbb{R}). \end{aligned}$$

[Exercise 33, Fall 1997]

Determine if it is better to persist with pointwise convergence of the first step, or to start dealing with L^2-convergence.
Note we are integrating the product of a Fourier series and a function having bounded variation in the third step of (5.1.h d).

Letting $\lambda = \gamma/2^m$ in the right side of (5.1.h d), we obtain

$$\text{(5.1.h e)} \qquad \begin{aligned} f_m(t) &= \sqrt{2\Omega}\sum_n c_{m,n}\int_{-2\Omega}^{2\Omega} \widehat{\psi}(\lambda) e^{2\pi i (2^m t - n/(4\Omega))\lambda} 2^m d\lambda \\ &= \sqrt{2\Omega}\sum_n c_{m,n} 2^m \psi(2^m t - \frac{n}{4\Omega}) \qquad \text{in } L^2(\mathbb{R}). \end{aligned}$$

Because of (5.1.h c) and the definitions of F_m and f_m on $\widehat{\mathbb{R}}$ and $\mathbb{R}$, resp., we have

$$f = \sqrt{2\Omega} f * \phi + \sum_{m=0}^{\infty} f_m = \sum f_m \quad \text{in } L^2(\mathbb{R});$$

and so we can apply (5.1.h e) to assert that

$$\begin{aligned} f(t) &= \sqrt{2\Omega} f * \phi(t) + \sum_{m=0}^{\infty} \sum_{n} d_{m,n} 2^{m/2} \psi(2^m t - \frac{n}{4\Omega}) \\ &= \sum_m \sum_n d_{m,n} 2^{m/2} \psi(2^m t - \frac{n}{4\Omega}) \end{aligned}$$

in $L^2(\mathbb{R})$, where $d_{m,n} = \sqrt{2\Omega} 2^{m/2} c_{m,n}$.

c. The Fourier transform of $f * d_{2\pi\Omega} \in PW_\Omega$ is $F_{-\infty} \equiv F\mathbf{1}_{[-\Omega,\Omega)}$. If we consider $F_{-\infty}$ as a 2Ω-periodic function defined on $\widehat{\mathbb{R}}$, then by the Fourier inversion formula for Fourier transforms we can compute, as in (5.1.h d), that

$$f * d_{2\pi\Omega}(t) = \int_{-\Omega}^{\Omega} F(\gamma) e^{2\pi i t \gamma} d\gamma \quad \text{in } L^2(\mathbb{R}),$$

and so

$$\begin{aligned} f * d_{2\pi\Omega}(t) &= \int_{-\Omega}^{\Omega} \sum_n F_{-\infty}{}^{\vee}[n] e^{-2\pi i n \gamma/(2\Omega)} e^{2\pi i t \gamma} d\gamma \\ \text{(5.1.h g)} \qquad & \\ &= \sum_n F_{-\infty}{}^{\vee}[n] \int_{-\Omega}^{\Omega} e^{2\pi i (t - \frac{n}{2\Omega}\gamma)} d\gamma \\ &= \sum_n F_{-\infty}{}^{\vee}[n] d_{2\pi\Omega}(t - \frac{n}{2\Omega}) \quad \text{in } L^2(\mathbb{R}), \end{aligned}$$

where

$$\begin{aligned} \forall n \in \mathbb{Z}, \qquad F_{-\infty}{}^{\vee}[n] &= \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\gamma) e^{2\pi i n \gamma/(2\Omega)} d\gamma \\ \text{(5.1.h h)} \qquad & \\ &= \frac{1}{2\Omega} f * d_{2\pi\Omega}(\frac{n}{2\Omega}) = \frac{1}{2\Omega} f * \phi(\frac{n}{2\Omega}). \end{aligned}$$

Thus,

$$\text{(5.1.h i)} \qquad f * \phi = \frac{1}{2\Omega} \sum f * \phi(\frac{n}{2\Omega}) \tau_{n/(2\Omega)} \phi,$$

and this is a form of the Classical Sampling Theorem for the function $f * \phi \in PW_\Omega$. □

Of course, $f * \phi = \frac{1}{\sqrt{2\Omega}} f$ in case $f \in PW_\Omega$. In this case (5.1.h i) is the Classical Smapling formula.

We have chosen the notation

$$F_{-\infty} = F\mathbf{1}_{[-\Omega,\Omega)}$$

since

$$F\mathbf{1}_{[-\Omega,\Omega)} = \sum_{m=-\infty}^{-1} F_m.$$

a. Eventually compare the Haar wavelet

$$\psi(t) = \begin{cases} 1 & \textbf{if } t \in [0, 1/2) \\ -1 & \textbf{if } t \in [1/2, 1) \\ 0 & \textbf{otherwise} \end{cases}$$

with the Shannon wavelet. There is a type of duality, which we will analyze, as well as the phenomenon of "extreme cases". Superficially, the Haar is rectangular and the Fourier transform of the Shannon wavelet is rectangular.

b. Besides doing "Shannon wavelet" decomposition to generalize beyond the Classical Sampling Theorem and simultaneously have a model for aliasing, we could also use the Gabor decomposition with its model for aliasing (by JB and Heller), see [W:M&A, Theorem 7.2.2]. - See ahead.

c. When we discuss MRAs, do a thorough job of tying together the Shannon wavelet decomposition with MRAs. In particular, let

$$V_0 = \overline{span}\left\{ \frac{1}{\sqrt{2\Omega}} \tau_{n/(2\Omega)} d_{2\pi\Omega} \right\}$$

and define the other V_js as in the definition of an MRA. As *Exercises*, show that $\cap V_j = \{0\}, \overline{\cup V_j} = L^2(\mathbb{R})$, etc., for the case of $\phi = \frac{1}{\sqrt{2\Omega}} d_{2\pi\Omega}$. In particular, do all of the details related to $\left\{\tau_{n/(2\Omega)}\phi\right\}$ as an ONB for $PW_\Omega = V_0$. Then obtain/observe that the first two equations of (5.1.h a) are the statement that

$$L^2(\mathbb{R}) = V_0 \oplus \bigoplus_{m=0}^{\infty} W_m$$

d. Reinforce the "wavelet condition" $\widehat{\psi}(0) = 0$ for the Shannon wavelet, from its origins as something like an "admissibility condition" for the reconstruction formula for continuous wavelet transforms. See Section 5.1.3 for Calderón formula and "need" to have $\widehat{\psi}(0) = 0$.

e. We choose the notation "$d_{m,n}$" to denote "detail". We shall be more precise about this in describing "debluring", see my vgs on Daubechies - Mallat for the explanation I should include in the following example.

Example 5.1.9 (Orthonormality Criteria for Shannon Wavelet Systems)

Let $\psi = \psi_{(\Omega)}$ be the Shannon wavelet.
Facts

a. $\{\psi_{m,n/(2\Omega)}\}$ **is ON.**

b. $\{\psi_{m,n/(4\Omega)}\}$ **is not orthogonal.**

(Since

$$a^{1/2}f(at-b) \leftrightarrow \frac{1}{a^{1/2}}e^{-2\pi i(\frac{b}{a})\gamma}F(\frac{\gamma}{a}), \qquad a>0$$

we have that

$$\begin{aligned} &\left\|2^{m/2}\psi(2^m t-b)\right\|^2_{L^2(\mathbb{R})} \\ = &\int \frac{1}{2^m}\left|\widehat{\psi}(\frac{\gamma}{2^m})\right|^2 d\gamma = \int \left|\widehat{\psi}(\lambda)\right|^2 d\lambda = 1. \qquad) \end{aligned}$$

Insert my notes on this material

Make figures of $\psi_{m,n/(4\Omega)}$ **as well as** $(\psi_{m,n/(4\Omega)})^\wedge$**.**

Theorem 5.1.10 (Shannon Wavelet Systems: ONBs and Tight Frames)

Let $\psi = \psi_{(\Omega)}$ be the Shannon wavelet.

a. $\{\psi_{m,n/(2\Omega)}\}$ is an ONB for $L^2(\mathbb{R})$.

b. $\{\psi_{m,n/(4\Omega)}\}$ is a tight frame for $L^2(\mathbb{R})$ with frame constants $A = B = 2$.

Make sure I define frames in Section 5.1, even though I don't develop them until Section 5.5.

Perhaps incorporate this result as part of "Orthonormality Criteria for Shannon Wavelet Systems" example.

State from my Chui or W:M&A7*. Do not prove it, but perhaps give as exercise or wait until frames in Section 5.5.
Mention that Gabor did it for Gaussians as we explained with "Gabor's Idea" section.

*** Theorem 7.22 which was proved by Heller & JB**

Note from W:M&A, Chapter 7:

The *Classical Sampling Theorem* and its generalizations , *Theorem 7.22* and *Theorem 7.24*, have several proofs, each ultimately depending on the relation between Fourier series and Fourier transforms. We shall give a frame theoretic proof of *Theorem 7.22* to illustrate a natural means of dealing with aliasing and with non-bandlimited sampling formulas, cf., [Ben 92, Section 2, Section 10, and Definition/Discussion 28] for aliasing and classical proofs, as well as [BS].
The sampling function $d_{2\pi\Omega}$ of *Theorem 7.21* is replaced by rather general sampling functions s in *Theorem 7.22*. These force oversampling but can give rise to faster convergence in applications. Although the hypotheses seem technical, the critical hypotheses are that $2T\Omega \leq 1$ and $\widehat{s} = 1$ on $[-\Omega, \Omega]$.

Theorem 5.1.11 (Gabor Decomposition (BH90))

Let T, $\Omega > 0$ be constants for which $0 < 2T\Omega \leq 1$, and let $g \in PW_{1/(2T)}$ have the properties that $\widehat{g} \in L^\infty(\widehat{\mathbb{R}})$,

$$\widehat{g} = 1 \qquad \text{on } [-\Omega, \Omega],$$

and, in case $2T\Omega < 1$, $\widehat{g}$ is continuous and

$$|\widehat{g}| > 0 \qquad \text{on } (\frac{-1}{2T}, -\Omega] \cup [\Omega, \frac{1}{2T}).$$

Set

$$G(\gamma) = \sum |\widehat{g}(\gamma - mb)|^2 \quad \text{and} \quad s(t) = (\frac{\widehat{g}}{G})^\vee(t),$$

where $\Omega + \frac{1}{2T} \leq b < \frac{1}{T}$ in case $2T\Omega < 1$ and $\Omega + \frac{1}{2T} = b$ if $2T\Omega = 1$. Then

$$\exists A, B > 0 \quad \text{such that} \quad A \leq G(\gamma) \leq B \quad \text{a.e.},$$

$s \in PW_{1/(2T)}$, $\widehat{s} \in L^\infty(\widehat{\mathbb{R}})$, $\widehat{s} = 1$ on $[-\Omega, \Omega]$,

$$\forall f \in L^2(\mathbb{R}), \quad f = T \sum \langle \widehat{f}, e_{nT}\tau_{mb}\widehat{g}\rangle \tau_{-nT}(e_{mb}s) \quad \text{in } L^2(\mathbb{R}),$$

and

$$\forall f \in PW_\Omega, \quad f = T \sum f(nT)\tau_{nT}s \quad \text{in } L^2(\mathbb{R}).$$

Technically, *Theorem 7.22* yields *Theorem 7.21* by making the special choice $b = 2\Omega$.

Example 5.1.12 (Smooth sampling functions.)

The functions $s \in PW_{1/(2T)}$ *of Theorem 7.22 can be chosen in the Schwartz space* $S(\mathbb{R})$ *(Definition 2.4.3) when* $2T\Omega < 1$. We begin the construction by defining

$$\psi_\epsilon(\gamma) = \frac{\varphi(\epsilon - |\gamma|^2)}{\int \varphi(\epsilon - |\lambda|^2)d\lambda},$$

where $\varphi \in C^\infty(\widehat{\mathbb{R}})$ vanishes on $(-\infty, 0]$ and equals $e^{-1/\gamma}$ on $[0, \infty)$. Thus $\psi_\epsilon \in C_c^\infty(\widehat{\mathbb{R}})$ is an even function satisfying the conditions, supp $\psi_\epsilon = [-\epsilon, \epsilon]$ and $\int \psi_\epsilon(\gamma)d\gamma = 1$. Next set

$$\psi_{U,V} = \frac{1}{||V||}\mathbf{1}_V * \mathbf{1}_{U-V}, \qquad U, V \subseteq \widehat{\mathbb{R}},$$

where $0 < |V| < \infty$ and U is compact. Thus, $\psi_{U,V}$ is 1 on U and vanishes off of $U + V - V$. The function $g \in S(\mathbb{R})$ of *Theorem 7.22* is defined in terms of $\widehat{g}$ as $\widehat{g} = \psi_{U,V} * \psi_\epsilon$, where we specify ϵ, U, and V as follows. Let $U = [-u, u]$, where $u \in (\Omega, \frac{1}{2T})$ is arbitrary, and let $\epsilon = u - \Omega$. Choose $V = [-v, v]$ by setting $v = \frac{w-u}{2}$, where $w = \frac{1}{2T} - \epsilon$. These choices are necessitated by a simple geometric argument, and the resulting function $\widehat{g}$ satisfies the conditions of *Theorem 7.22* and $\widehat{g} \in C_c^\infty(\widehat{\mathbb{R}})$. The definition of s in *Theorem 7.22* completes the construction.

Remark 5.1.13 (The Classical Sampling Theorem and Aliasing)

Recall Section 3.10, which will be part of HAA4

Insert remarks about decomposing in terms of a sequence of functions with bounded bandwidth (with the Shannon Wavelet Decomposition). Compare with Gabor case and the way in which the aliased part can be decomposed. Refer back to page 258.

Use my aliasing material from
- my Chui Chapter
- W:M&A7, Chapter 7.
Comment on Aliasing pseudo - measures.

Problem area: Tie-in the "aliased" part of a general Gabor expansion, i.e., $\sum_{m\neq 0, n\in\mathbb{Z}}$, with the conventional explanation of aliasing in terms of overlapped Fourier series, i.e., in the case $2T\Omega > 1$.

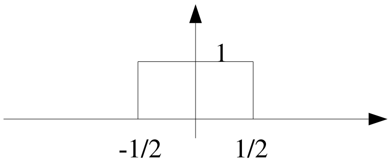


Figure 9: f

5.2 Multiresolution Analysis Wavelet ONBs

Definition 5.2.1 (Orthonormal Wavelets)

Let $\psi \in L^2(\mathbb{R})$. For each $m, n \in \mathbb{Z}$, $\psi_{m,n}$ is the function defined by

$$\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n) \tag{4}$$

The function ψ is an **orthonormal wavelet** if $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.

Mention Shannon and Haar again.
Exercise Prove $\nexists$ON wavelet ψ s.t. $\psi \in C_c^\infty(\mathbb{R})$.
Modify this exercise. Recall Meyer's original example in frequency domain, where modulates and dilates of some special $\psi \in C_c^\infty(\mathbb{R})$ give $\{\psi_{m,n}\}$ as an ONB.

Theorem (Daubechies 1987)
Let $r \geq 1$. $\exists\psi \in C_c^{(r)}(\mathbb{R})$, such that $\{\psi_{m,n}\} \subseteq L^2(\mathbb{R})$ is an ONB for $L^2(\mathbb{R})$.
where $\{\psi_{m,n}\}$ is the **affine system**. $\{\psi_{m,n}\}$ defined by

$$\forall m, n \in \mathbb{Z}, \psi_{m,n}(t) = 2^{m/2}\psi(2^m - n).$$

Example 5.2.2 (Fourier and Wavelet Expansions)

A form of the Haar wavelet was defined in Remark1.6.4b.

After defining the Haar orthonormal wavelet ψ precisely, compare the Fourier and wavelet coefficients of the function f (see Figure 9) with its perturbation g (see Figure 10).

Note that the inner product $< g, \psi_{m,n} >= 0$ when g is constant on the non-zero set of $\psi_{m,n}$, see Figure 11 and Figure 12.

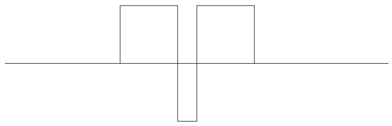


Figure 10: g

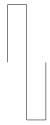


Figure 11: $\psi_{m,n}$

Quote Exercise 3.14 where I reference [Dau92]. Use figure of above function as I do with my viewgraphs.
Addend Exercise 3.14, referring ahead to §5.4.
Addend Remark 1.6.4b, referring ahead to §5.4.

Thus, the wavelet coefficients of a signal f and its perturbation g won't change except in the region of perturbation.

If f and g are 1-periodic, note that

$$\widehat{f}[n] = \begin{cases} 0 & : \quad n \neq 0 \\ 1 & : \quad n = 0 \end{cases}$$

and that, in general, $\hat{g}$ will never be 0.

Thus, small adjustments to f give rise to a change of **all** of its frequency information.

Hence we do **not** have **frequency localization**.

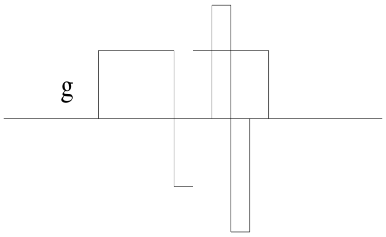


Figure 12:

These inner products are not frequency information, but do provide **scale localization**. In some cases, there is a relationship between scale and frequency, e.g., [BC95] on ECoG analysis.

Problem area Use this relationship to quantify any relationship between frequency and scale localization.

Besides **scale localization**, wavelet decompositions have the feature of **progressive resolution** ("deblurring"), which, in turn, allows for the effective implementation of thresholding for compression and minimization of cost problems.

> **See $5.2.c for progressive resolution.**

Thus small disturbances in a signal generally will affect all values of the Fourier transform.

Looked at slightly differently, we can say that an inherent liability of the Fourier transform is that its "basis functions" on $\mathbb{R}$, viz.,

$$\{e_{-\gamma} : \gamma \in \widehat{\mathbb{R}}\},$$

are supported on the whole real line. Thus, information about a function f is required on all of $\mathbb{R}$ in order to obtain any value of the Fourier transform $\widehat{f}$. Compactly supported orthonormal wavelet bases avoid this disadvantage and provide scale localization.

Example 5.2.3 (Deblurring and Multiresolution)

> **Put in my "deblurring" vg, and insert my spiral figure explaining multiresolution.**
> **In these figures, only do the outward spiral, noting the inner can also be made.**
> **What type of multiresolution would a 3-dimensional spiral designate?**

Let ψ be an ON wavelet. The concept of **multiresolution** is motivated and understood by the following nomenclature and intuition.

If $f \in L^2(\mathbb{R})$ then $f = \sum_{m,n} < f, \psi_{m,n} > \psi_{m,n}$.

Suppose $supp\ \psi \subseteq [-1/2, 1/2]$, thus,

$$supp\ \psi_{m,n} \subseteq I_{m,n},$$

where

$$I_{m,n} = [\frac{n}{2^m} - \frac{1}{2^{m+1}}, \frac{n}{2^m} + \frac{1}{2^{m+1}}].$$

Thus, the length of $I_{m,n}$ is $|I_{m,n}| \equiv 2^{-m}$. Let

$$f_M = \sum_{m \leq M} \sum_{n} < f, \psi_{m,n} > \psi_{m,n}.$$

Consequently, $f_{M+1} = f_M + \sum_n < f, \psi_{M+1,n} > \psi_{M+1,n}$ can be thought of as **deblurring** f_M by adding to f_M the behavior of f on intervals of length $2^{-(M+1)}$.

Definition 5.2.4 (Multiresolution Analysis(MRA))

An MRA of $L^2(\mathbb{R})$ is a sequence $\{V_j : j \in \mathbb{Z}\}$ of closed subspaces of $L^2(\mathbb{R})$ satisfying the following properties:

a. $\forall j \in \mathbb{Z}, V_j \subseteq V_{j+1}$;

b. $f \in V_0$ if and only if $\tau_k f \in V_0$ for all $k \in \mathbb{Z}$;

c. $f(t) \in V_j$ if and only if $f(2t) \in V_{j+1}$;

d. $\bigcap V_j = \{0\}$ and $\overline{\bigcup V_j} = L^2(\mathbb{R})$;

e. $\exists \phi \in V_0$ such that $\{\tau_k \phi\}$ is an ONB for V_0.

The function ϕ is a **scaling function** for the $\{V_j\}$.

> **Shannon and Haar orthonormal wavelets are derived from MRAs.**
> **a. Do Shannon case as exercise(essentially #34 Fall 1997).**
> **b. We shall prove the Haar assertion in** §5.2.8.
> **By "subspace" we mean "linear" or "vector" subspace.**

The following theorem will be proved in Section 5.6. The rationale for the proof will be given in Section 5.2.6.

Theorem 5.2.5 (Mallat and Meyer Theorem)

Let $\{V_j : j \in \mathbb{Z}\}$ be an MRA of $L^2(\mathbb{R})$ with scaling funciton ϕ. There is a constructible function ψ, depending on ϕ such that $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$. In fact, a choice for the function ψ is

$$\psi(t) = \sqrt{2} \sum_n h_1[n] \phi(2t - n) \ in L^2(\mathbb{R}),$$

where

$$\forall n \in \mathbb{Z}, \ h_1[n] = (-1)^n \overline{h_0[-n+1]},$$

and where $\{h_0[n]\}$ is the sequence of Fourier coefficients of the QMF $H_0 \in L^\infty(\mathbb{T})$, which in turn is a solution of the frequency scaling equation

$$\sqrt{2}\widehat{\phi}(2\gamma) = H_0(\gamma)\widehat{\phi}(\gamma) \ in L^2(\mathbb{R});$$

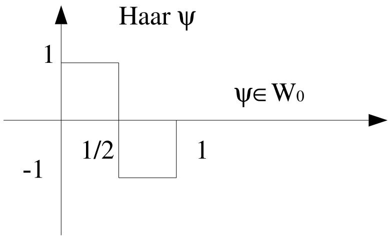


Figure 13: **Haar function ψ**

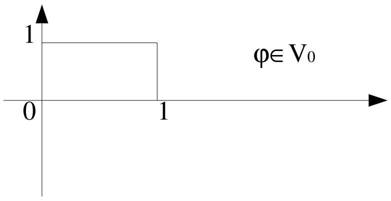


Figure 14: **Haar scaling function ϕ**

Define a quadrature mirror filter(QMF) H, see [Ben92, page501]
Note $\|H\|_{L^2(\mathbb{T})} = 1$ and $\|H\|_{L^\infty(\mathbb{T})} \leq \sqrt{2}$.
Do Perfect Reconstruction by QMFs-see [Ben92] and my proofs thereof.
[Mal89]

a. There are QMFs that are not derived from MRAs. This depends on Cohen's use of congruence criteria, e.g., Kluvanek, Lagarias, Yang Wang, Gröchenig, Madych, ...

b. Larson and Dai have constructed orthonormal wavelets ψ for $L^2(\mathbb{R}^d)$ for $d \geq 2$. In particular, one can conclude in this case that the ONB $\{\psi_{m,n} : m \in \mathbb{Z}, n \in \mathbb{Z}^d\}$ is not derived from an MRA of Ł$^2(\mathbb{R}^d)$. This was mentioned in a Note in Section 5.1 also.

Theorem 5.2.6 (Outline: MRA Algorithm) *MRA Algorithm to Construct ONBs for $L^2(\mathbb{R})$, an interplay between Fourier Transforms and Fourier Series*

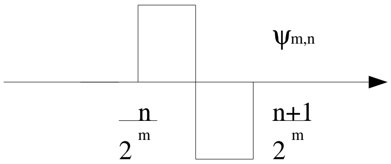


Figure 15: **Haar functions** $\psi_{m,n}$

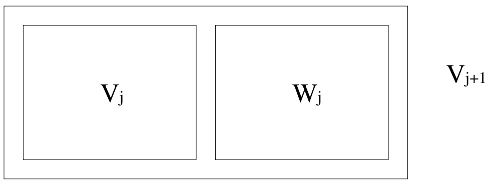


Figure 16:

Let $\{V_j : j \in \mathbb{Z}\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function ϕ,

a. Let W_j be the **orthogonal complement** of V_j in V_{j+1}(*B.*). Thus, W_j, is a closed subspace of V_{j+1}, and V_{j+1} has the direct sum decomposition $V_{j+1} = V_j \oplus W_j$, cf., Exercise 5.2.7. Recall that this direct sum means that each element $f \in V_{j+1}$ has the unique representation $f = f_V + f_W$, where $f_V \in V_j$ and $f_W \in W_j$. All of the direct sums we shall consider in Section 5.2 are, in fact, orthogonal complement direct sums, e.g., in the direct sum above, we have $W_j = V_j^{\perp}$ ans so $V_{j+1} = V_j \oplus V_j^{\perp}$ where

$$V_j^{\perp} = \{f \in V_{j+1} : \forall g \in V_j, < f, g >= 0\}.$$

Give a reference for this fact, e.g., [GG81 ,page21] or prove it or put it in AppendixB.

By property **iii** of an MRA

(A)

$$f(t) \in W_j \ if \ and \ only \ if \ f(2t) \in W_{j+1}.$$

The proof of (A) goes as follows. Let $f \in W_m$. Then $f \in V_{m+1}$ since $V_{m+1} = V_m \oplus W_m$. Thus, $f(t/2) \in V_m = V_{m-1} \oplus W_{m-1}$. On the other hand, $f(t/z) \perp V_{m-1}$ since

$$2^{(m-1)/2} \int \phi(2^{m-1}t - n) \frac{1}{\sqrt{2}} \overline{f(t/2)} dt = \int 2^m \phi(2^m u - n) \overline{f(u)} du,$$

and the right side vanishes since $f \in W_m$, $\phi_{m,n} \in V_m$, and $V_{m+1} = V_m \oplus W_m$, see **Figure 17**.

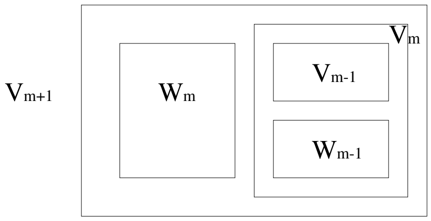


Figure 17:

Since $f(t/2) \perp V_{m-1}$ and $f(t/2) \in V_m$ we conclude that $f(t/2) \in W_{m-1}$, and hence, (A) is valid.

b. A key feature of the MRA algorithm is the construction of a function $\psi \in W_0$ for which $\{\tau_k \psi\}$ is an ONB for W_0. Suppose such a function ψ has been constructed.

c. Because of (A), we can prove that

(B)

$$\forall j \in \mathbb{Z},\ fixed,\ \{\psi_{j,k} : k \in \mathbb{Z}\}$$

is an ONB for W_j. We shall prove this in Section 5...

d. Consider the spaces V_j and W_j, $j \in \mathbb{Z}$. Figure 16 exhibits the direct sum relationship $V_{j+1} = V_j \oplus W_j$. Figure 18 exhibits the direct sum relationship for several values of j. The geometric content of this relationship is that the sequence $\{V_j\}$ is increasing, i.e.,

$$\forall j \in \mathbb{Z},\ V_j \subseteq V_{j+1},$$

and that the sequence $\{W_j\}$ is disjoint, except for 0 which is in each W_j.

e. The sequences $\{V_j\}$ and $\{W_j\}$ both encompass all of $L^2(\mathbb{R})$, but in different ways. For the sequence $\{V_j\}$, property **iv** of an MRA asserts that $\overline{\bigcup V_j} = L^2(\mathbb{R})$. For the "disjoint" sequence $\{W_j\}$, the decomposition of $L^2(\mathbb{R})$ is denoted by

(C)

$$L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j.$$

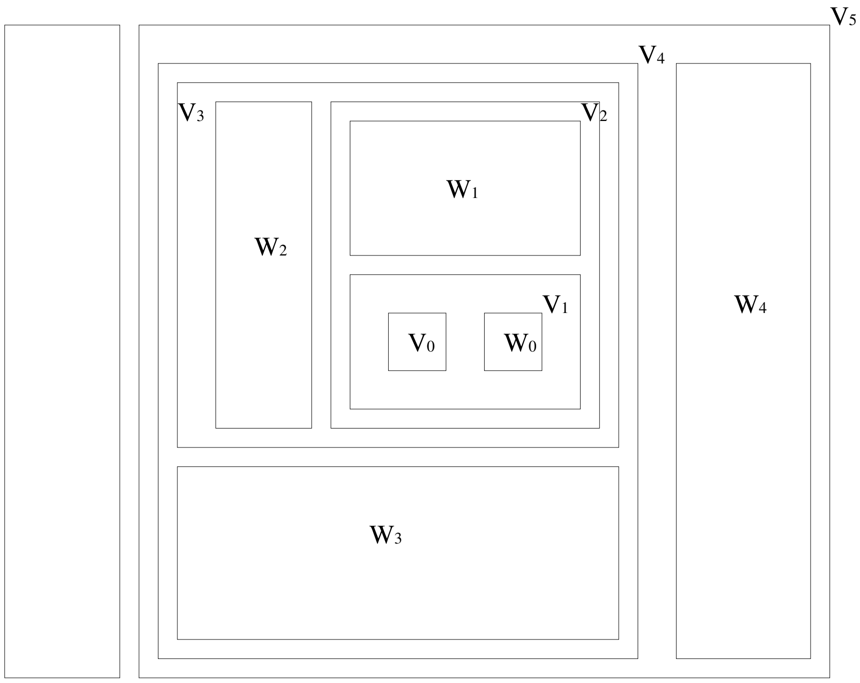


Figure 18:

The direct sum notation on the right side of (C) will be made precise in Section 5... For the purpose of this outline, it suffices to say that (C) combines with (B) to allow us to prove that

(D)

$$\{\psi_{j,k} : j, k \in \mathbb{Z}\} \textit{ is an ONB for } L^2(\mathbb{R}).$$

f. (D) is the main assertion of the MRA algorithm. The main step in the algorithm required for this assertion is the construction of a function $\psi \in W_0$ such that $\{\tau_n \psi\}$ is an ONB for W_0, see part b. We shall now show how this construction is made.

g. Notationally,

$$\phi_{-1,0} = \frac{1}{\sqrt{2}}\phi(\frac{t}{2})$$

Using properties **iii** and **v** of an MRA, we have

$$\phi_{-1,0} \in V_{-1} \subseteq V_0,$$

and hence,

(E)

$$\frac{1}{\sqrt{2}}\phi(\frac{t}{2}) = \sum_n h_0[n]\phi(t-n) \textit{ in } L^2(\mathbb{R}),$$

where $\{h_0[n]\} \in \ell^2(\mathbb{R})$ is defined by

(F)

$$\forall n \in \mathbb{Z},\ h_0[n] = \int \phi_{-1,0}(t)\overline{\phi(t-n)}dt.$$

(The subscript "0" in the coefficient sequence $\{h_0[n]\}$ will be explained in Section 5...)

| **Quote relevant Theorem c. ONB-decomposition from App B(2nd)**

h. Note that a function $f \in L^2(\mathbb{R})$ is an element of W_{-1} if and only if $f \in V_0$ and $f \perp V_{-1}$, cf., the proof of (A). Thus the desired function ψ must have the property that $\psi_{-1,0} \in V_0$ and $\psi_{-1,0} \perp V_{-1}$, where $\psi_{-1,0}(t) \equiv \frac{1}{\sqrt{2}}\psi(t/2)$.

Since $\psi_{-1,0} \in V_0$ and since $\{\tau_n \phi\}$ is an ONB for V_0, then ψ, **assuming** it exists, has the property that its dilation $\psi_{-1,0}$ can be expressed as

(G)

$$\frac{1}{\sqrt{2}}\psi(\frac{t}{2}) = \sum_n h_1[n]\phi(t-n) \textit{ in } L^2(\mathbb{R}),$$

where $\{h_1[n]\} \in \ell^2(\mathbb{Z})$ is defined by

(H)

$$\forall n \in \mathbb{Z},\ h_1[n] = \int \psi_{-1,0}(t)\overline{\phi(t-n)}dt.$$

Theorem B.X(HAA(2nd)) again

Taking the Fourier transform of both sides of equations (E) and (G) gives

(I)

$$\sqrt{2}\widehat{\phi}(2\gamma) = H_0(\gamma)\widehat{\phi}(\gamma)\ in\ \mathrm{L}^2(\widehat{\mathbb{R}})$$

and

(J)

$$\sqrt{2}\widehat{\psi}(2\gamma) = H_1(\gamma)\widehat{\phi}(\gamma)\ in\ \mathrm{L}^2(\widehat{\mathbb{R}}),$$

respectively, where

$$H_0(\gamma) = \sum_n h_0[n]e^{-2\pi in\gamma}\ and\ H_1(\gamma) = \sum_n h_1[n]e^{-2\pi in\gamma}.$$

Rescaling, we write (J) as

(K)

$$\widehat{\psi}(\gamma) = H_1(\frac{\gamma}{2})\frac{1}{\sqrt{2}}\widehat{\phi}(\frac{\gamma}{2})\ in\ \mathrm{L}^2(\widehat{\mathbb{R}}).$$

Note. Explain "in $\mathrm{L}^2(\widehat{\mathbb{R}})$" in (I), (J), and (K).

The annihilator constraint $\psi_{-1,0} \perp V_{-1}$ combines with (K) through some calculations to yield the fact that ψ, defined by the Fourier transform relation

(L)

$$\widehat{\psi}(\gamma) = -e^{-\pi i\gamma}\overline{H_0(\frac{\gamma}{2}+\frac{1}{2})}\frac{1}{\sqrt{2}}\widehat{\phi}(\frac{\gamma}{2})\ in\ L^2(\widehat{\mathbb{R}}),$$

is an element of W_0, see Theorem 5.7... for details.

i. For this function ψ, further calculations, which are not difficult and which are also provided in Theorem 5.7..., allow us to conclude that $\{\tau_n\psi\}$ is an ONB for W_0.

j. Because of (K) and (L) and the structure of ϕ and H_0, we are able to deduce that

(M)

$$H_1(\gamma) = -e^{-2\pi i\gamma}\overline{H_0(\gamma+1/2)},$$

see Theorem 5.7 for details. Thus,

(N)

$$\forall n \in \mathbb{Z},\ h_1[n] = (-1)^n \overline{h_0[-n+1]}.$$

In fact,

$$\begin{aligned} h_1[n] &= \int_{\mathbb{T}} H_1(\gamma) e^{2\pi i n\gamma}\, d\gamma = -\int_{\mathbb{T}} \overline{H_0(\gamma + 1/2)} e^{2\pi i (n-1)\gamma} d\gamma \\ &= -\int_{\mathbb{T}} \overline{H_0(\lambda)} e^{2\pi i(n-1)\lambda} e^{-\pi i(n-1)} d\lambda \\ &= (-1)^n \overline{h_0[-n+1]}. \end{aligned}$$

k. Summing up, for a given MRA and corresponding scaling function ϕ, the MRA algorithm is the following sequence of steps:

   i. Construction of H_0 by means of equation(F);

   ii. Definition of ψ by means of (L) and the inverse Fourier transform;

   iii. The computation of $\{\psi_{m,n}\}$, which, in turn is an ONB for $L^2(\mathbb{R})$.

   We reiterate that in this outline we have omitted verifications of our assertions. These will come in Section 5.7... Further, we have invoked a time honored method for defining the function ψ. We have shown **if** the desired function ψ exists, then its Fourier transform has the form asserted in (L), cf., Section1.8 where the same strategy is used. Once we have equation (L) we can prove that ψ defined by (L) gives rise to an ONB for $L^2(\mathbb{R})$.

   Finally, the orthonormal wavelet for a given MRA and scaling function is not uniquely determined.

   | **Reference ahead, etc.**

Example 5.2.7 (The Haar MRA and ONB)

a. Let $\varphi = \mathbf{1}_{[0,1)}$ and let $V_0 \equiv \overline{\text{span}}\{\tau_n \varphi\}$. Then $\{\tau_n \varphi\}$ is an ONB for V_0, e.g., *Theorem B.*

| **Note: ONB≡ON&Complete**

Define
$$V_j = \{f \in L^2(\mathbb{R}) : 2^{j/2} f(2^j t) \in V_0\}.$$
It is easy to check that each V_j is a closed subspace of $L^2(\mathbb{R})$. It can be proved that $\{V_j\}$ is an MRA of $L^2(\mathbb{R})$ with scaling function φ, e.g., *Proposition 5.2.8.* This comes down to proving the following not unreasonable facts:
$$\begin{array}{ll} i. & \forall j \in \mathbb{Z},\ \ V_j \subseteq V_{j+1}; \\ ii. & \bigcap V_j = \{0\}; \\ ii. & \overline{\bigcup V_j} = L^2(\mathbb{R}). \end{array}$$

b. In this example we shall assume that $\{V_j\}$ is an MRA with scaling function φ, and then construct an orthonormal wavelet from the MRA algorithm.

Since $\varphi = \mathbf{1}_{[0,1)}$, we compute
$$\begin{aligned} h_0[n] &= \frac{1}{\sqrt{2}} \int \varphi(\frac{t}{2}) \mathbf{1}_{[0,1)}(t-n)dt \\ &= \frac{1}{\sqrt{2}} \int \mathbf{1}_{[0,1)}(\frac{t}{2}) \mathbf{1}_{[0,1)}(t-n)dt = \frac{1}{\sqrt{2}} \int_0^2 \mathbf{1}_{[0,1)}(t-n)dt, \end{aligned}$$
see *Figure X.* Thus,
$$h_0[n] = \begin{cases} \frac{1}{\sqrt{2}} & n = 0,\ 1, \\ 0 & \text{otherwise.} \end{cases}$$
Therefore, by the MRA algorithm we obtain
$$h_1[n] = (-1)^n \overline{h_0[-n+1]};$$
and so
$$\begin{aligned} h_1[0] = &\quad \overline{h_0[1]} &= \frac{1}{\sqrt{2}} \\ h_0[1] = & -\overline{h_0[0]} &= -\frac{1}{\sqrt{2}} \end{aligned}$$

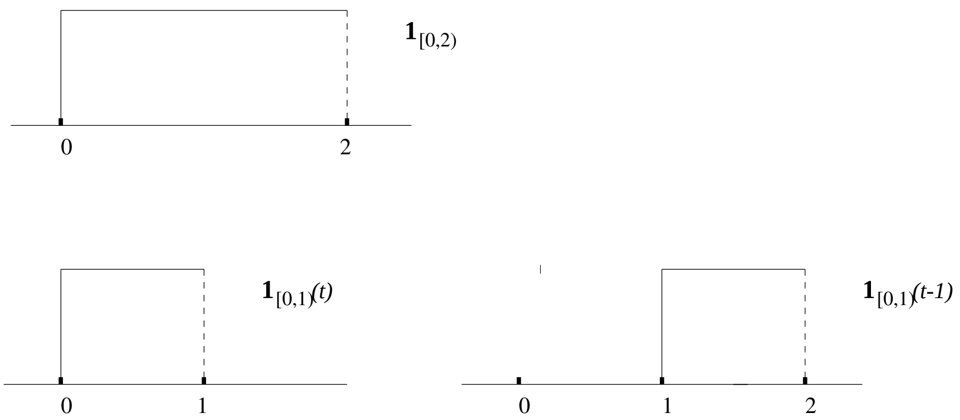


Figure 19: Figure X

and

$$h_1[n] = 0 \quad \text{if } n \neq 0,\ 1.$$

Further, the MRA algorithm allows us to write

$$\psi(t) = \sum h_1[n]\sqrt{2}\varphi(2t-n);$$

and we compute

$$\varphi(2t-n) = \begin{cases} 1 & 2t-n \in [0,1), \text{ i.e., if } t \in [\frac{n}{2}, \frac{n+1}{2}) \\ 0 & \text{otherwise,} \end{cases}$$

so that

$$\varphi(2t-0) = \begin{cases} 1 & \text{if } t \in [0, \frac{1}{2}) \\ 0 & \text{otherwise} \end{cases}$$

and

$$\varphi(2t-1) = \begin{cases} 1 & \text{if } t \in [\frac{1}{2}, 1) \\ 0 & \text{otherwise.} \end{cases}$$

Combining this with the definition of h_1, we see that a wavelet ψ corresponding to the given MRA with scaling function $\varphi = \mathbf{1}_{[0,1)}$ is

(5) $$\psi = \mathbf{1}_{[0,1/2)} - \tau_{1/2}\mathbf{1}_{[0,1/2)},$$

see *Figure A*.

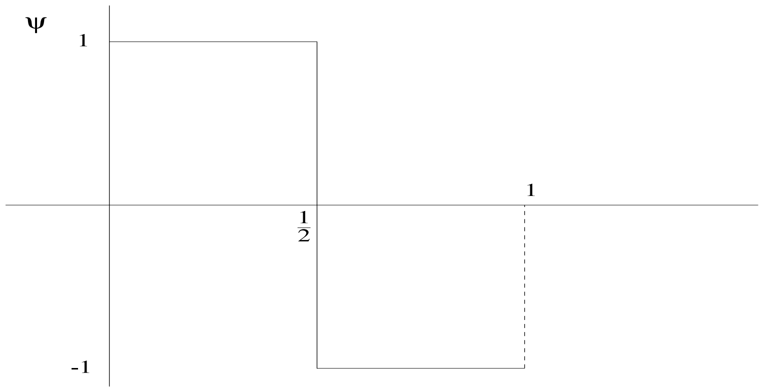


Figure 20: Figure A

As such, we refer to the sequence $\{V_j\}$ of *part a* as the **Haar MRA** with scaling function φ. The function ψ in (5) is the **Haar wavelet**.

Now, let's prove that the sequence $\{V_j\}$ and function φ of *Example 5.2.7* is an MRA. We shall provide details for the case of $L^p(\mathbb{R})$, $p \in [1, \infty)$.

Proposition 5.2.8 (Haar MRA for $L^p(\mathbb{R})$, $p \in [0, \infty)$)
Let $\varphi = \mathbf{1}_{[0,1)}$, let $p \in [1, \infty)$, and set

$$V_0 = \{\sum a_n \tau_n \varphi : \{a_n\} \in l^p(\mathbb{Z})\},$$

For each $j \in \mathbb{Z}$, define

$$V_j = \{g(t) = 2^{j/p} f(2^j t) : f \in V_0\}.$$

Then $\{V_j\}$ is an MRA of $L^p(\mathbb{R})$ with scaling function φ.

Proof: It is easy to show that each V_j is a closed subspace of $L^p(\mathbb{R})$.

i. We shall prove

$$\forall j \in \mathbb{Z}, \qquad V_j \subseteq V_{j+1},$$

$$\bigcap V_j = \{0\},$$

and

$$\overline{\bigcup V_j} = L^p(\mathbb{R}).$$

ii. Note that

$$\forall m \in \mathbb{Z}, \quad V_m = \{\sum_n a_n \mathbf{1}_{[\frac{n}{2^m}, \frac{n+1}{2^m})} : \sum |a_n|^p < \infty\} \subseteq L^p(\mathbb{R}).$$

In fact

$$\begin{aligned} & \int |\sum_n c_n \mathbf{1}_{[\frac{n}{2^m}, \frac{n+1}{2^m})}(t)|^p \, dt \\ = & \sum_n \int_{\frac{n}{2^m}}^{\frac{n+1}{2^m}} |c_n|^p \, dt = \frac{1}{2^m} \sum |c_n|^p. \end{aligned}$$

iii. Let $f \in V_m$. Then

$$f = \sum a_n \mathbf{1}_{[\frac{n}{2^m}, \frac{n+1}{2^m})} = \sum b_k \mathbf{1}_{[\frac{k}{2^{m+1}}, \frac{k+1}{2^m})}, \tag{6}$$

where $\sum_n |a_n|^p < \infty$ and where $b_k = a_{[\frac{k}{2}]}$, i.e., $b_{2n} = a_n$ and $b_{2n+1} = a_n$. To see that $f \in V_{m+1}$ we must show that $f = \sum c_n \mathbf{1}_{[\frac{n}{2^{m+1}}, \frac{n+1}{2^{m+1}})}$, where $\sum |c_n|^p < \infty$. By (6), this comes down to proving

$$\sum |b_k|^p < \infty,$$

which is true since $\sum |b_k|^p = 2 \sum |a_n|^p < \infty$.

iv. Assume $f \in V_j$ for all $j \in \mathbb{Z}$. $f \in V_0$ means that $f = \sum a_n \mathbf{1}_{[n,n+1)}$, where $\sum |a_n|^p < \infty$. Since $f \in V_{-1}$, then the graph in *Figure A* can not occur. In fact, $f = \sum b_n \mathbf{1}_{[\frac{n}{2^{-1}}, \frac{n+1}{2^{-1}})}$ implies $b_0 = a_0$ and $b_0 = a_1$, etc., i.e.,

$$a_0 = a_1, \ a_2 = a_3, \ldots$$

and

$$a_{-1} = a_{-2}, \ a_{-3} = a_{-4}, \ldots$$

cf., *Figure B.*

Generally, $f \in V_{-m}$, $m \geq 1$, implies

$$a_0 = a_1 = \ldots = a_{2^m - 1}$$

and

$$a_{-1} = a_{-2} = \ldots = a_{-2^m}.$$

Since $\sum |a_n|^p < \infty$ we conclude that each $a_n = 0$, i.e., $f = 0$. Hence, $\bigcap V_j = \{0\}$.

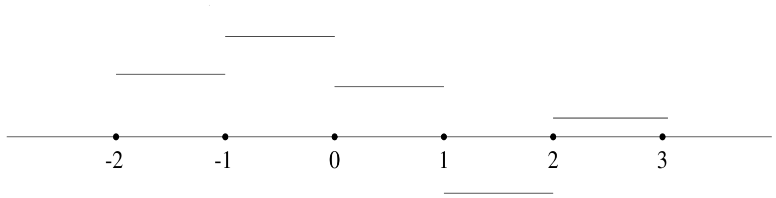


Figure 21: Figure A

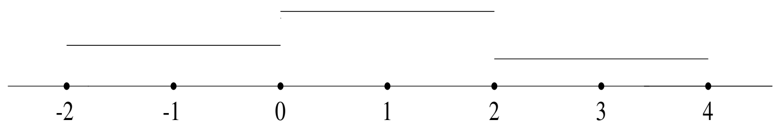


Figure 22: Figure B

v. Finally, we shall prove that $\overline{\bigcup V_j} = L^p(\mathbb{R})$. Let $f \in L^p(\mathbb{R})$ and let $\epsilon > 0$. There is a step function $g = \sum_{k=1}^{N} c_k \mathbf{1}_{[a_k, b_k)}$, $b_k \leq a_{k+1}$, for which

$$\|f - g\|_{L^p(\mathbb{R})} < \epsilon/2 \tag{7}$$

| **Appendix A.5a and** *Exercise 5.31.*

Let $C = \max_{k=1,\cdots,N} |c_k|$, and choose m so that

$$\frac{1}{2^m} < \frac{(\epsilon/2)^p}{2NC^p}. \tag{8}$$

We shall construct an element $h \in V_m$ so that $\|f - h\|_{L^p(\mathbb{R})} < \epsilon$; and this yields the desired density. For each k, we approximate $[a_k, b_k)$ from inside by intervals of the form $I_{m,k} = [\frac{r_k}{2^m}, \frac{s_k}{2^m})$, where r_k, $s_k \in \mathbb{Z}$. If $b_k - a_k < \frac{2}{2^m}$ we choose the empty set. The r_k and s_k are further chosen so that

$$\frac{r_k - 1}{2^m} < a_k \leq \frac{r_k}{2^m} \quad \text{and} \quad \frac{s_k}{2^m} \leq b_k < \frac{s_k + 1}{2^m}.$$

The sequence $\{I_{m,k} : k = 1, \cdots, M\}$ is disjoint, and

$$\forall k = 1, \cdots, N, \ |[a_k, b_k) \backslash I_{m,k}| < \frac{2}{2^m}.$$

Let $h = \sum_{k=1}^{N} c_k \mathbf{1}_{I_{m.k}}$, and note that $h \in V_m$. Further,

$$\begin{aligned} ||g-h||^p_{L^p(\mathbb{R})} &= \int |\sum_{k=1}^{N} c_k(\mathbf{1}_{[a_k,b_k)}(t) - \mathbf{1}_{I_{m,k}}(t))|^p dt \\ &= \sum_{k=1}^{N} |c_k|^p((b_k - \frac{s_k}{2^m}) + (\frac{r_k}{2^m} - a_k)) \qquad (9) \\ &\leq \frac{2}{2^m}\sum_{k=1}^{N}|c_k|^p \leq \frac{2NC^p}{2^m} < (\frac{\epsilon}{2})^p, \end{aligned}$$

where the last step follows by (8).

Combining inequalities (7) and (9) with*Minkowski's Inequality* (*Remark A.17a*), we obtain

$$||f-h||_{L^p(\mathbb{R})} \leq ||f-g||_{L^p(\mathbb{R})} + ||g-h||_{L^p(\mathbb{R})} < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon,$$

which is the desired result since $h \in V_m$. The use of *Minkowski's Inequality* requires $p \geq 1$. □

It is sometimes possible to verify the MRA density condition $\overline{\bigcup V_j} = L^1(\mathbb{R})$ for more complicated scaling function than $\varphi = \mathbf{1}_{[0,1)}$. This can be done by means of *Wiener's Tauberian Theorem*, e.g., *Section 2.9.12* and *2.9.13* as well as [Ben75], [BZ97].

Example 5.2.9 (MRA and Change of Basis)

Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\varphi \in V_0$ and orthonormal wavelet $\psi \in W_0$, constructed as in *Section 5.2.6*. Thus, $\{\tau_n\varphi\}$ is an ONB for V_0 and $\{\tau_n\varphi\}$ is an ONB for W_0.

a. Besides $\{\tau_n\varphi\}$, *the sequence of functions*

$$\frac{1}{\sqrt{2}}\varphi(\frac{t}{2}-n),\ \frac{1}{\sqrt{2}}\psi(\frac{t}{2}-n), \qquad (10)$$

where $n \in \mathbb{Z}$, *forms an ONB for* V_0.

To see this, first note that

$$\forall n \in \mathbb{Z}, \quad \frac{1}{\sqrt{2}}\varphi(\frac{t}{2}-n) \in V_{-1}. \qquad (11)$$

In fact, if $f(t) \equiv \varphi(t-n)$, then $f \in V_0$ so that $f(\frac{t}{2}) \in V_{-1}$ by definition of an MRA. Thus, (11) is obtained since $f(\frac{t}{2}) = \varphi(\frac{t}{2}-n)$ and since V_{-1} is a linear subspace of $L^2(\mathbb{R})$.

Next, we can assert that

$$\forall n \in \mathbb{Z}, \quad \frac{1}{\sqrt{2}}\psi(\frac{t}{2} - n) \in W_{-1}. \tag{12}$$

In fact, if $g(t) = \psi(t-n)$, then $g \in W_0$ so that $g(\frac{t}{2}) \in W_{-1}$ (as we noted in *Section 5.2.6a*). Thus (12) is obtained since $g(\frac{t}{2}) = \psi(\frac{t}{2} - n)$ and since W_{-1} is a vector subspace of $L^2(\mathbb{R})$.

Exercise: Prove that for every fixed m, $\{\varphi_{m,n} : n \in \mathbb{Z}\}$ is an ONB for V_m(for a given MRA). (*Exercise 5.8*)

The orthonormality of the functions in (10) is a consequence of the orthonormality of $\{\tau_n\varphi\}$ in V_0, the orthonormality of $\{\tau_n\psi\}$ in W_0, and the fact that $V_0 = V_{-1} \oplus W_{-1}$, an orthonormal complement direct sum. For example,

$$\int \frac{1}{\sqrt{2}}\varphi(\frac{t}{2} - n)\frac{1}{\sqrt{2}}\overline{\varphi(\frac{t}{2} - q)}dt = \int \varphi(u-n)\overline{\varphi(u-q)}du = \delta(n,q).$$

Finally, the fact that the functions in (10) form an ONB for V_0 is a consequence of the just proved orthonormality and the following calculations for arbitrary $f \in V_0$:

$$f = f_V + f_W, \text{ where } f_V \in V_{-1},\ f_W \in W_{-1};$$

$f_V(2t) \in V_0$ so that

$$f_V(2t) = \sum \langle f_V(2u), \tau_n\varphi(u)\rangle \tau_n\varphi(t) \text{ in } L^2(\mathbb{R}),$$

and thus,

$$f_V(t) = \sum \langle \sqrt{2} f_V(2u), \tau_n\varphi(u)\rangle \frac{1}{\sqrt{2}}\varphi(\frac{t}{2} - n) \in V_{-1};$$

and finally $f_W(2t) \in W_0$ so that

$$f_W(2t) = \sum \langle f_W(2u), \tau_n\psi(u)\rangle \tau_n\psi(t) \text{ in } L^2(\mathbb{R}),$$

and thus

$$f_W(t) = \sum \langle \sqrt{2} f_W(2u), \tau_n\psi(u)\rangle \frac{1}{\sqrt{2}}\psi(\frac{t}{2} - n) \in W_{-1},$$

e.g., *Theorem B*.

ONB≡ON+Completeness. INSERT IN APP. B

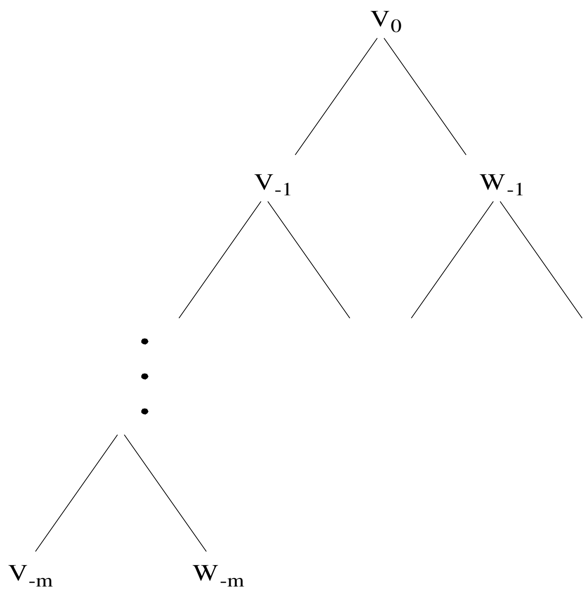


Figure 23: Figure Y

b. The procedure of *part a* can be used to make other ONBs for V_0. In fact, for any fixed $m \in \mathbb{N}$ we have the orthogonal complement direct sum

$$V_0 = V_{-m} \oplus W_{-m} \oplus W_{-m+1} \oplus W_{-m+2} \oplus \ldots \oplus W_{-1};$$

and hence the sequence of functions

$$\varphi_{-m,n}\,,\ \psi_{-m,n}\,,\ \psi_{-m+1,n}\,,\cdots,\psi_{-1,n}\,,$$

where $n \in \mathbb{Z}$, forms an ONB for V_0, see *Figure Y*.

This procedure can be generalized significantly by introducing the notion of wavelet packets.

Outline 5.2.10 (Wavelet Packet Algorithm)

Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\varphi \in V_0$ and orthonormal wavelet $\psi \in W_0$, constructed as in *Section 5.2.6*.

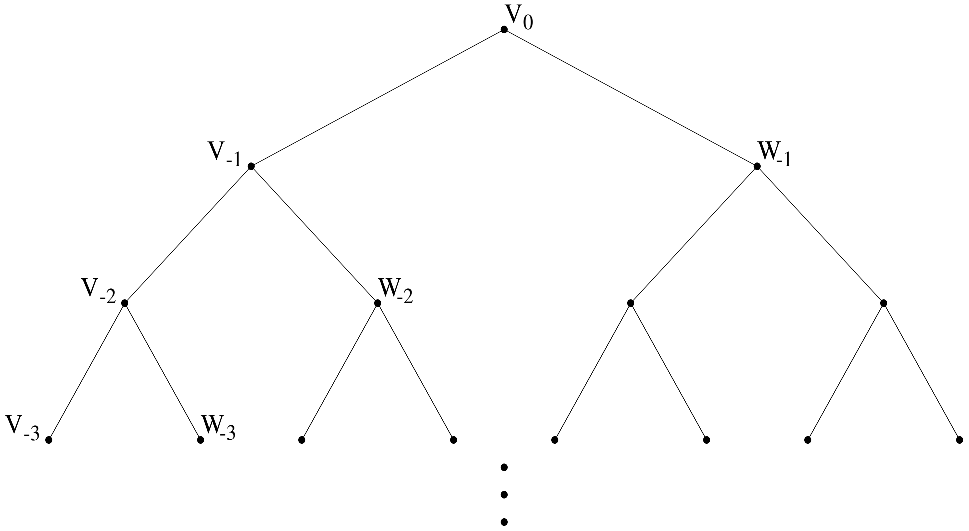


Figure 24: Figure A

From the MRA algorithm we have

$$\varphi(t) = \sum h_0[n]\varphi_{1,n}(t) \in V_1 \textbf{ in } L^2(\mathbb{R})$$

and

$$\psi(t) = \sum h_1[n]\varphi_{1,n}(t) \in V_1 \textbf{ in } L^2(\mathbb{R}).$$

| **State L^2 convergence more conveniently in displayed items.**

Thus,

$$\widehat{\varphi}(\gamma) = \frac{1}{\sqrt{2}}\widehat{\varphi}(\frac{\gamma}{2})H_0(\frac{\gamma}{2}) \textbf{ and } \widehat{\psi}(\gamma) = \frac{1}{\sqrt{2}}\widehat{\varphi}(\frac{\gamma}{2})H_1(\frac{\gamma}{2}) \textbf{ in } L^2(\widehat{\mathbb{R}}).$$

| **Carefully explain what "in $L^2(\mathbb{R})$" means.**

The ONB splitting of V_0 for a given MRA, reflected by (10),

| **of** *Section 5.2.9*

can be effected at every node in the dyadic tree of *Figure A*. (At this point we have only labeled the nodes V_0, V_{-1}, W_{-1}, V_{-2}, W_{-2}, V_{-3}, W_{-3}, $\cdots$.)

To see how this is done, we begin with the ONB $\{\tau_n \varphi\}$ for V_0, or, equivalently, the ONB $\{e_{-n}\widehat{\varphi}\}$ for $\widehat{V}_0 \subseteq L^2(\widehat{\mathbb{R}})$. Then the nodes designated V_{-1} and W_{-1} can be characterized by the formulas written in *Figure B*.

$$\begin{array}{cc} V_{-1} & W_{-1} \\ \varphi_{-1,n}(t) = \frac{1}{\sqrt{2}}\varphi(\frac{t}{2}-n) \longleftrightarrow e_{-2n}(\gamma)\sqrt{2}\widehat{\varphi}(2\gamma), & \psi_{-1,n}(t) = \frac{1}{\sqrt{2}}\psi(\frac{t}{2}-n) \longleftrightarrow e_{-2n}(\gamma)\sqrt{2}\widehat{\psi}(2\gamma) \\ \sqrt{2}\widehat{\varphi}(2\gamma) = \widehat{\varphi}(\gamma)H_0(\gamma) & \sqrt{2}\widehat{\psi}(2\gamma) = \widehat{\varphi}(\gamma)H_1(\gamma) \end{array}$$

Figure B

In particular, $\{e_{-2n}(\gamma)\sqrt{2}\widehat{\varphi}(2\gamma)\}$ is an ONB for $\widehat{V}_{-1}$ and $\left\{e_{-2n}(\gamma)\sqrt{2}\widehat{\psi}(2\gamma)\right\}$ is an ONB for $\widehat{W}_{-1}$.

As indicated in *Section 5.2.9*, we obtain an ONB splitting of V_{-1} by the sequence

$$\{(\frac{1}{\sqrt{2}})^2\varphi(\frac{t}{2^2}-n),\ (\frac{1}{\sqrt{2}})^2\psi(\frac{t}{2^2}-n) :\ n \in \mathbb{Z}\}, \tag{13}$$

noting that

$$\{(\frac{1}{\sqrt{2}})^2\varphi(\frac{t}{2^2}-n)\} \subseteq V_{-2} \text{ and } \{(\frac{1}{\sqrt{2}})^2\psi(\frac{t}{2^2}-n)\} \subseteq W_{-2}. \tag{14}$$

For example, if $f \in V_{-2}$ and $g(t) \equiv f(2t)$, then $g \in V_{-1}$ so that

$$g(t) = \sum \langle g, \varphi_{-1,n}\rangle \frac{1}{\sqrt{2}}\varphi(\frac{t}{2}-n) \text{ in } L^2(\mathbb{R});$$

and hence

$$f(t) = \sum \langle \sqrt{2}g, \varphi_{-1,n}\rangle (\frac{1}{\sqrt{2}})^2\varphi(\frac{t}{4}-n) \text{ in } L^2(\mathbb{R}).$$

In the case of (13) and (14), the analogue of *Figure B*, for the nodes designated V_{-2} and W_{-2} coming from the splitting of V_{-1}, is given by *Figure C*.

$$\begin{array}{cc} V_{-2} & W_{-2} \\ \varphi_{-2,n}(t) = (\frac{1}{\sqrt{2}})^2\varphi(\frac{t}{4}-n) & \psi_{-2,n}(t) = (\frac{1}{\sqrt{2}})^2\psi(\frac{t}{4}-n) \\ \longleftrightarrow & \longleftrightarrow \\ e_{-4n}(\gamma)(\sqrt{2})^2\widehat{\varphi}(4\gamma), & e_{-4n}(\gamma)(\sqrt{2})^2\widehat{\psi}(4\gamma), \\ (\sqrt{2})^2\widehat{\varphi}(4\gamma) = \widehat{\varphi}(\gamma)H_0(\gamma)H_0(2\gamma) & (\sqrt{2})^2\widehat{\psi}(4\gamma) = \widehat{\varphi}(\gamma)H_0(\gamma)H_1(2\gamma) \end{array}$$

Figure C

In particular, $\{e_{-4n}(\gamma)(\sqrt{2})^2\widehat{\varphi}(4\gamma)\}$ is an ONB for $\widehat{V}_{-2}$ and $\{e_{-4n}(\gamma)(\sqrt{2})^2\widehat{\psi}(4\gamma)\}$ is an ONB for $\widehat{W}_{-2}$. Clearly, the last formulas in *Figure C* are consequences of equation (13), e.g.,

$$(\sqrt{2})^2\widehat{\varphi}(4\gamma) = \sqrt{2}\widehat{\varphi}(2\gamma)H_0(2\gamma) = \widehat{\varphi}(\gamma)H_0(\gamma)H_0(2\gamma).$$

At this point it is reasonable to ask if there is a natural splitting of W_{-1} in *Figure A* to complete the remaining two nodes(in *Figure A*) which are on the same level as V_{-2} and W_{-2}.

The answer is "yes" and the procedure results in new ONBs for V_0.

| **It should be pointed out that the result is not unique, e.g., [BS94], [BLS98].**

| **[BS94] is Benedetto-Saliani SPIE and [BLS98] is Benedetto-Leon-Saliani.**

In order to describe the procedure it is convenient to use the notation of *Remark 3.9.4e*. Thus,

| *Remark 3.9.4e* **will be in new HAA4**

at level 0, X_0^0 denotes V_0; at level 1, X_0^1 and X_1^1 denote V_{-1} and W_{-1}; at level 2, $X_{(0,0)}^2$ and $X_{(0,1)}^2$ denote V_{-2} and W_{-2}.

In the case of level 1, the subscript 0 of X_0^1 corresponds to the subscript 0 of H_0 on the left side of *Figure B*, and the subscript 1 of X_1^1 corresponds to the subscript 1 of H_1 on the right side of *Figure B*.

In the case of level 2, there are the four consecutive spaces $X_{(0,0)}^2$, $X_{(0,1)}^2$, $X_{(1,0)}^2$, and $X_{(1,1)}^2$ from *Remark 3.9.4*. The subscript $(0,0)$ of $X_{(0,0)}^2 = V_{-2}$ corresponds to the two subscripts 0 and 0 of $H_0(\gamma)$ and $H_0(2\gamma)$ on the left side of *Figure C*. The subscript $(0,1)$ of $X_{(0,1)}^2 = W_{-2}$ corresponds to the two subscripts 0 and 1 of $H_0(\gamma)$ and $H_1(2\gamma)$ on the right side of *Figure C*.

Because of this pattern we associate the function $\widehat{\varphi}(\gamma)H_1(\gamma)H_0(2\gamma)$ with $X_{(1,0)}^2$ and the function $\widehat{\varphi}(\gamma)H_1(\gamma)H_1(2\gamma)$ with $X_{(1,1)}^2$. In this way, level 2 of the tree in *Figure A* is completed. Thus, we define the functions $\theta_{1,0}$ and $\theta_{1,1}$ by the formulas

$$(\sqrt{2})^2\widehat{\theta}_{1,0}(2^2\gamma) = \widehat{\varphi}(\gamma)H_1(\gamma)H_0(2\gamma)$$

and

$$(\sqrt{2})^2\widehat{\theta}_{1,1}(2^2\gamma) = \widehat{\varphi}(\gamma)H_1(\gamma)H_1(2\gamma).$$

Then $X_{(1,0)}^2$ is defined as

$$X_{(1,0)}^2 = \overline{\mathrm{span}}\{\frac{1}{(\sqrt{2})^2}\theta_{1,0}(\frac{t}{2^2} - n) : n \in \mathbb{Z}\} \text{ in } L^2(\mathbb{R})$$

and $X^2_{(1,1)}$ is defined as

$$X^2_{(1,1)} = \overline{\text{span}}\{\frac{1}{(\sqrt{2})^2}\theta_{1,1}(\frac{t}{2^2} - n) : n \in \mathbb{Z}\} \text{ in } L^2(\mathbb{R}).$$

It is easy to prove that $\{\frac{1}{(\sqrt{2})^2}\theta_{1,0}(\frac{t}{2^2} - n)\}$ is an ONB for $X^2_{(1,0)}$, that $\{\frac{1}{(\sqrt{2})^2}\theta_{1,1}(\frac{t}{2^2} - n)\}$ is an ONB for $X^2_{(1,1)}$, and that

$$W_{-1} = X^1_1 = X^2_{(1,0)} \oplus X^2_{(1,1)},$$

an orthogonal complement direct sum.

The previous details give rise to the following general procedure for a given MRA. Using the notation from *Remark 3.9.4*, we form the tree $\{X^r_m : r \geq 1\}$ where $m \in \{0, 1, 2, 3, \cdots, 2^r - 1\}$ and where $X^0_0 \equiv V_0$. Each element X^r_m of the tree is a closed linear subspace of V_0. Further, X^r_m, $r \geq 1$, is determined by the function

$$(\frac{1}{\sqrt{2}})^r \widehat{\theta}_{\epsilon_1,\cdots,\epsilon_r}(2^r\gamma) = \widehat{\varphi}(\gamma)H_{\epsilon_1}(\gamma)H_{\epsilon_2}(2\gamma)H_{\epsilon_3}(2^2\gamma)\ldots H_{\epsilon_r}(2^{r-1}\gamma), \tag{15}$$

$m = \sum_{j=1}^r \epsilon_j 2^{j-1}$ and $\epsilon_j \in [0, 1]$, in the sense that

$$X^r_m = \overline{\text{span}}\{\frac{1}{(\sqrt{2})^r}\theta_{\epsilon_1,\cdots,\epsilon_r}(\frac{t}{2^r} - n) : n \in \mathbb{Z}\}, \tag{16}$$

$$\{\frac{1}{(\sqrt{2})^r}\theta_{\epsilon_1,\cdots,\epsilon_r}(\frac{t}{2^r} - n) : n \in \mathbb{Z}\} \text{ is an ONB for } X^r_m, \tag{17}$$

and

$$V_0 = \bigoplus_{m=0}^{2^r-1} X^r_m, \tag{18}$$

an orthogonal complement direct sum.

Verify these direct sums are orthog. complements, and so state. This is but one orthogonal complement direct sum decomposition of V_0 at level r. When should I discuss hemlines? (boundaries of trees?)

The **Wavelet Packet Algorithm** is the tree $\{X^r_m\}$ of subspaces of V_0 defined by (15), (16), (17), and (18).

MODIFY THIS DEFINITION

Although we have written the orthogonal complement direct sum in (18) in terms of the natural ordering $0, 1, \cdots, 2^r - 1$, recall from *Remark 3.9.4* that at level r the spaces X^r_m are ordered, left to right on the tree, by bit reversal ordering.

Eventually use products, such as in tree, to define scaling functions. Tie in with Viète formula.

Definition 5.2.11 (Wavelet Packet Filter)

Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\varphi \in V_0$ and orthonormal wavelet $\psi \in W_0$. Let H_0 be the QMF associated with φ and let H_1 be the QMF associated with ψ. The sequence of **wavelet packet filters** corresponding to this structure is defined by the sequence of products,

$$\prod_{j=1}^{r} H_{\epsilon_j}(2^j\gamma),$$

where $r \geq 1$, $m \in \{0, 1, \ldots, 2^r - 1\}$, and $\epsilon_j \in \{0, 1\}$.

Does $\exists WP$ splitting for FMRAs that is productive? See [JFAA3, 4].

Q. When should I do Walshes, sequency, etc? A. *Section 5.3*
When should I state W-P theorems? (the advanced theory) See Ryan translation for clear statements. Perhaps state a Theorem clearly before drawings pics.

-Incorporate MATLAB W-P exercises and perhaps one example for text.
-Make mathematical exercises w W-Ps.

Example 5.2.12 (Wavelet Packet ONBs and Frequency Analysis)

Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$. In terms of subspaces and change of bases, we can illustrate the inclusions and orthogonal complements of the $\{V_j\}$ and $\{W_j\}$ by means of Figure 5.30.

MRA Structure

Figure 5.30

Besides this MRA structure, an MRA gives rise to a decomposition of the frequency band $[-\frac{1}{2}, \frac{1}{2})$ associated with the QMFs H_0, H_1 and their dilates.

To fix ideas suppose $H_0 \equiv \sqrt{2}\mathbf{1}_{[-\frac{1}{4},\frac{1}{4})}$. Then H_0 is a QMF and $\widehat{\varphi} = 1_{[-\frac{1}{2},\frac{1}{2})}$ is a solution of the scaling equation

$$\sqrt{2}\widehat{\varphi}(2\gamma) = \widehat{\varphi}(\gamma)H_0(\gamma).$$

If we set $V_0 = \overline{span}\{\tau_n\varphi\}$ and define

$$V_j = \{g(t) = 2^{\frac{j}{2}}f(2^jt) : f \in V_0\},$$

then $\{V_j\}$ is an MRA of $L^2(\mathbb{R})$ with scaling function φ, and the orthonormal wavelet corresponding to $\{V_j\}$ is the Shannon wavelet $\psi = \psi_{(\frac{1}{2})}$.

Figure 5.31 illustrates the aforementioned frequency decomposition in the case $H_0 = \sqrt{2}\mathbf{1}_{[-\frac{1}{4},\frac{1}{4})}$.

MRA Frequency Structure

Figure 5.31

In the case of wavelet packets, the pyramidal subspaces and change of basis structure, described in Section 5.2.10, is illustrated in Figure 5.33. As with MRAs, there is a corresponding frequency decomposition, which we have illustrated in Figure 5.34 in terms of the Shannon MRA.

Figure 5.32

MRA Wavelet Packet Structure

Figure 5.33

MRA Wavelet Packet Frequency Structure

Figure 5.34

Figure 5.35

Example 5.2.13 (2-dimensional MRAs)

a. Let ψ be an orthonormal wavelet for $L^2(\mathbb{R})$. For each $(j,k) \in \mathbb{Z}^2$ and $(m,n) \in \mathbb{Z}^2$, define the tensor product

$$\psi_{j,k} \otimes \psi_{m,n}(x_1, x_2) = \psi_{j,k}(x_1)\psi_{m,n}(x_2).$$

Then $\{\psi_{j,k} \otimes \psi_{m,n} : j,k,m,n \in \mathbb{Z}\}$ is an ONB for $L^2(\mathbb{R}^2)$. As such we shall say that $\psi \otimes \psi$ is a rectilinear orthonormal wavelet for $L^2(\mathbb{R}^2)$ if for every ordered 4-tuple $(j,k,m,n,) \in \mathbb{Z}^4$, we define

$$(\psi \otimes \psi)_{j,k,m,n}(x_1, x_2) = \psi_{j,k} \otimes \psi_{m,n}(x_1, x_2).$$

(The term "rectilinear" will be explained shortly.)

Notice that $(\psi \otimes \psi)_{j,k,m,n}$ provides separate dilation and translation in the x_1 and x_2 variable. For example, if $supp\, \psi = [-\frac{1}{2}, \frac{1}{2}]$ then $supp\, \psi_{j,k} = [\frac{k}{2^j} - \frac{1}{2^{j+1}}, \frac{k}{2^j} + \frac{1}{2^{j+1}}]$. Thus the support of $(\psi \otimes \psi)_{j,k,m,n}$ is the rectangle in Figure 5.36.

Figure 5.36

Consequently, the wavelet decomposition of arbitrary $f \in L^2(\mathbb{R}^2)$ invloves rectangles of all shapes, e.g., every long and skinny, in order to achieve reconstruction. We refer to this approach as the *rectilinear tensor product wavelet decomposition.*

b. Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function φ and orthonormal wavelet ψ constructed from the MRA algorithm. Define the tensor product

$$\varphi \otimes \varphi(x_1, x_2) = \varphi(x_1)\varphi(x_2),$$

and let V_0^2 be the set of functions $f : \mathbb{R}^2 \to \mathbb{C}$ having the from

$$f(x_1, x_2) = \sum_{n_1, n_2 \in \mathbb{Z}} c_{n_1,n_2} \varphi \otimes \varphi(x_1 - n_1, x_2 - n_2),$$

where $c = \{c_{n_1,n_2}\} \in l^2(\mathbb{Z}^2)$. Then, $V_0^2 = \overline{span}\{f \otimes g(x_1, x_2) : f, g \in V_0\} = V_0 \hat{\otimes} V_0$ (Exercise 5). Next, set

$$V_j^2 = \{g(x_1, x_2) = f(2^j x_1, 2^j x_2) : f \in V_0^2\}.$$

It can now be proved that $\{V_j^2 : j \in \mathbb{Z}\}$ is an MRA of $L^2(\mathbb{R}^2)$ with scaling function $\varphi \otimes \varphi$.

The orthogonal complement of V_0^2 in V_1^2 is denoted by W_0^2.

We were able to describe the MRA Structure and the MRA Wavelet Packet Structure of $L^2(\mathbb{R})$ in Figures ... and ... by means of a sequence of one-dimensional decompositions into spaces, ordered according to level. We can do the same for the MRA and MRA Wavelet Packet Structure of $L^2(\mathbb{R}^2)$ by a sequence of squares.

For example, if we start with $V_1^2 = V_0^2 \oplus W_0^2$, then the first level of the MRA structure is pictured in Figure 5.37.

Figure 5.37

At the next step, for the case of the MRA Structure of $L^2(\mathbb{R}^2)$, we conduct the same spliting, but this time on V_0^2, see Figure 5.38.

Figure 5.38

In these figures, the tensor products are defined as expected,viz.,

$$\varphi \otimes \psi(x_1, x_2) = \varphi(x_1)\psi(x_2),$$
$$\psi \otimes \varphi(x_1, x_2) = \psi(x_1)\varphi(x_2),$$
$$\psi \otimes \psi(x_1, x_2) = \psi(x_1)\psi(x_2).$$

This splitting is a natural consequence of the fact that

$$\begin{aligned} V_{j+1}^2 &= V_{j+1}\hat{\otimes}V_{j+1} = (V_j \oplus W_j)\hat{\otimes}(V_j \oplus W_j) \\ &= V_j\hat{\otimes}V_j \oplus [(W_j\hat{\otimes}V_j) \oplus (V_j\hat{\otimes}W_j) \oplus (W_j\hat{\otimes}W_j)]. \end{aligned}$$

c. Single Wavelets in $\mathbb{R}^d$.

Theorem Dai, Larson, Speegle Theorem [DLS]

There exists $K \subseteq \hat{\mathbb{R}}^d$ such that $\hat{\psi} = \mathbf{1}_K$ defines an orthonormal wavelet for $L^2(\mathbb{R}^d)$, i.e.,

$$\{\psi_{j,k}(x) = 2^{\frac{dj}{2}}\psi(2^j x - k) : j \in \mathbb{Z}, k \in \mathbb{Z}^d\}$$

is an ONB for $L^2(\mathbb{R}^d)$.

It was elementary to construct wavelet ONBs for $L^2(\mathbb{R}^d)$ of the form

$$\psi_{j_1,k_1} \otimes ... \otimes \psi_{j_d,k_d},$$

where ψ is an orthonormal wavelet for $L^2(\mathbb{R})$ and $j_1, ..., j_d, k_1, ..., k_d \in \mathbb{Z}$. The Dai, Larson, Speegle Theorem is a fundamentally deeper result since the dilation depends only on $j \in \mathbb{Z}$. Similarly, it is a deeper result than the MRA approach to $\mathbb{R}^d$.

<u>Definition</u> A set $K \subseteq \widehat{\mathbb{R}}^d$ is a *wavelet set* if the function ψ defined by $\hat{\psi} = \mathbf{1}_K$ is an orthonormal wavelet for $L^2(\mathbb{R}^d)$. A *partition* of a measurable set $A \subseteq \mathbb{R}^d$ is a sequence $\{A_n\}$ of measurable subsets $A_n \subseteq A$ such that $\bigcup A_n = A$ a.e. and $|A_m \cap A_n| = 0$ if $m \neq n$.

The following is true.

<u>Theorem</u> Dai, Larson, Speegle Theorem [DLS]

A measurable set $K \subseteq \mathbb{R}^d$ is a wavelet set if and only if

$$i.\ \{k + K : k \in \mathbb{Z}^d\} \text{ is a partition of } \mathbb{R}^d$$

and

$$ii.\ \{2^j K : j \in \mathbb{Z}\} \text{ is a partition of } \mathbb{R}^d.$$

<u>Remark</u> Condition *i* of the Theorem implies that $|K| = 1$. Also condition *i* suggests that K is some sort of a "cube", whereas condition *ii* suggests that K is some sort of an "annulus".

<u>Example</u> Wavelet set K

Let $K_1 = [-\frac{1}{2}, \frac{1}{2}]^2$, see Figure 5.39

Figure 5.39

To form K_2, first translate the cube $[0, \frac{1}{4}]^2$ by the vector (-1,-1)$\in \mathbb{Z}^2$. Thus translate

$$\begin{aligned} (-1,1) &: 1 \longrightarrow \text{3rd quadrant} \\ (1,-1) &: 2 \longrightarrow \text{4th quadrant} \\ (1,1) &: 3 \longrightarrow \text{1st quadrant} \\ (-1,-1) &: 4 \longrightarrow \text{2nd quadrant} \end{aligned}$$

We use these transformation since we can only use $\mathbf{Z}^d$-translations.

The shaded region of Figure 5.40 is K_2.

Next notice that $2^{-1}K_2 \cap K_2 \neq \phi$, e.g., $[\frac{3}{8}, \frac{1}{2}]^2 \subseteq 2^{-1}K_2 \cap K_2$. Similarly,$[\frac{3}{8}, \frac{1}{2}] \times [-\frac{3}{8}, -\frac{1}{2}] \subseteq 2^{-1}K_2 \cap K_2$, etc., see Figure 5.41. Part *ii* of the previous Theorem is the point of this calculation. Translate these smaller squares, as in the previous step, to obtain K_3. K_3 is the "annular extension" of the shaded region in Figure 5.42.

d. Single wavelets for $\mathbb{R}^d$ and MRA

The construction of part c reduces to the Shannon wavelet in $\mathbb{R}^1$; and the Shannon wavelet ψ is an MRA wavelet. Recall $\hat{\psi} = \mathbf{1}_{[-1,1)} - \mathbf{1}_{[-\frac{1}{2},\frac{1}{2})}$.

There are orthonormal wavelets in $L^2(\mathbb{R})$ which can not be derived from the MRA algorithm. The Journé wavelet ψ defined by $\hat{\psi} = \mathbf{1}_K$, where

$$K = [-\frac{16}{7}, -2) \cup [-\frac{1}{2}, -\frac{2}{7}) \cup [\frac{2}{7}, \frac{1}{2}) \cup [2, \frac{16}{7}),$$

is such an example.

Pascal Auscher [Aus 95] proved that if $\psi^1, .., \psi^M \in L^2(\mathbb{R}^d)$, if their Fourier transforms satisfy mild smoothness and decay conditions, and if

$$\{\psi^i_{j,k}(x) = 2^{\frac{dj}{2}} \psi^i(2^j x - k) : i = 1, ...M, j \in \mathbb{Z}, k \in \mathbb{Z}^d\}$$

is an ONB for $L^2(\mathbb{R}^d)$, then the set $\{\psi^i : i = 1, ...M\}$ comes from an MRA.

We shall say that $\{\psi^i : i = 1, ..., M\} \subseteq L^2(\mathbb{R}^d)$ is a *wavelet collection* if $\{\psi^i_{j,k}\}$ is an ONB for $L^2(\mathbb{R}^d)$. Auscher, Gripenberg, and Wang proved the following theorem.

<u>Theorem</u> Let $\psi_1, ..., \psi_M \in L^2(\mathbb{R}^d)$ be a wavelet collection. This collection is associated with an MRA having a single scaling function if and only if

$$D(\gamma) = \sum_{i=1}^{M} \sum_{j=1}^{\infty} \sum_{k \in \mathbb{Z}^d} |\hat{\psi}^i(2^j(\gamma + k)|^2 = 1 \quad \text{a.e.}.$$

Auscher proved that generally D is integer valued; and by Plancherel's Theorem we have

$$(*) \quad \int_{\mathbb{T}^d} D(\gamma) d\gamma = \frac{M}{2^d - 1}.$$

By $(*)$ and the previous theorem we can conclude that *if* $\psi_1, ..., \psi_M \in L^2(\mathbb{R}^d)$ *is a wavelet collection associated with an* MRA, *then* $M = 2^d - 1$.

Notice this was the case for $\mathbb{R}^d$ when we required the three (2^2-1) wavelets $\varphi\otimes\psi, \psi\otimes\varphi,$ and $\psi\otimes\psi$.

5.3 The Haar Wavelet System and Walsh Wavelet Packets

Definition 5.3.1 (Dyadic Intervals)
Let $m, n \in \mathbb{Z}$. The closed *dyadic interval* $I_{m,n} \subseteq \mathbb{R}$ is the closed interval $[\frac{n}{2^m}, \frac{n+1}{2^m}]$.

> **- Put this definition earlier in HAA(5).**
> **- Write out this definition for $\mathbb{R}^d$.**
> **- Use the odds and ends from Higgins' book on Haar, Rademachers, and Walshes; quote [Hig76].**
> **- Tie in with groups and dyadic expansions.**

We have already defined the Haar wavelet system $\{\psi_{mn}\}$, where the *Haar function* is defined as

$$\psi(t) = \begin{cases} 1 & \text{if} \quad t \in [0, 1/2), \\ -1 & \text{if} \quad t \in [1/2, 1), \\ 0 & \text{if} \quad \text{otherwise} \end{cases}$$

and where $\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n)$, cf., Remark 1.6.4, as well as [quote other uses of Haar in HAA(5)].
Let $m \in \mathbb{Z}$ and let $n \in \mathbb{Z}^d$. The closed *dyadic cube* $Q_{m,n} \subseteq \mathbb{R}^d$ is the closed set

$$\{x \in \mathbb{R}^d : 2^{-m} n_j \leq x_j \leq 2^{-m}(n_j + 1), j = 1, \ldots, d\}.$$

Let $\mathrm{Q} = \{\mathrm{Q}_{m,n} : m \in \mathbb{Z}, n \in \mathbb{R}^d\}$ and for $Q \in \mathrm{Q}$ let $l(Q)$ be the side length of Q.

> **Perhaps put more material here.**

Proposition 5.3.2 (Nesting Property)
Let $I_{m,n}$ and $I_{p,q}$ be two closed dyadic intervals. If the interiors of $I_{m,n}$ and $I_{p,q}$ have nonempty intersection, then either $I_{m,n} \subseteq I_{p,q}$ or $I_{p,q} \subseteq I_{m,n}$.

Proof. Assuming

$$\left(\frac{n}{2^m}, \frac{n+1}{2^m}\right) \cap \left(\frac{q}{2^p}, \frac{q+1}{2^p}\right) \neq \emptyset,$$

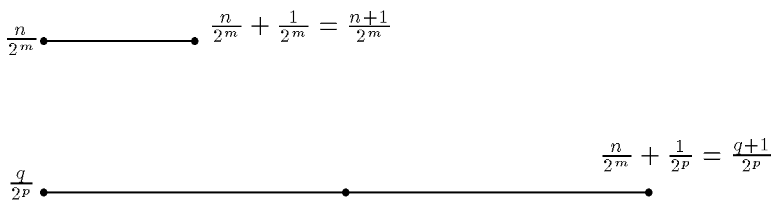
we can suppose, without loss of generality that

$$\frac{n}{2^m} \leq \frac{q}{2^p} < \frac{n+1}{2^m} \tag{19}$$

For case 1, assume $m = p$. Then (19) implies $n \leq q < n+1$ and so $n = q$. Then $I_{m,n} = I_{p,q}$.
For case 2, assume $m < p$ so that $2^{p-m} > 2$. Then (19) implies

$$\frac{n}{2^m} = \frac{n2^{p-m}}{2^p} \leq \frac{q}{2^p} < \frac{(n+1)2^{p-m}}{2^p} = \frac{n+1}{2^m}$$

So that $q + 1 \leq n2^{p-m}2^p$. Thus, $I_{p,q} \subseteq I_{m,n}$.
For case 3, assume $p < m$ and consider the two subcases, viz., $n/2^m = q/2^p$ and $n/2^m < q/2^p$.
Suppose $n/2^m = q/2^p$. Since $p < m$ we have $2^{m-p} \geq 2$, and so $1/2^m \leq 1/2^{p+1} < 1/2^p$. Thus, $I_{m,n} \subseteq I_{p,q}$, see Figure 5.3



Finally, suppose $n/2^m < q/2^p$ for the case $p < m$. Then, the distance d between the two points is

$$0 < d = \frac{q}{2^p} - \frac{n}{2^m} = \frac{q2^{m-p} - n}{2^m} = \frac{k}{2^m}$$

for some $k \in \mathbb{N}$. Thus, $(n+1)/2^m \leq q/2^p$, and this cannot occur by our assumption (19). □

Proposition 5.3.3 (Orthonormality of the Haar System)
Let ψ be the Haar function.

a. $\forall m, n \in \mathbb{Z}$, supp $\psi_{m,n} = I_{m,n}$.

b. The Haar wavelet system $\{\psi_{m,n}\}$ is othonormal.

Proof: Part a is clear. b.

i. Let $m = p$ and let $n = q$. Then

$$\|\psi_{m,n}\|^2_{L^2(\mathbb{R})} = < \psi_{m,n}, \psi_{p,q} > = 2^m \int_{\frac{n}{2^m}}^{\frac{n+1}{2^m}} 1 dt = 1.$$

ii. Let $m = p$ and let $n \neq q$. Then

$$[\frac{n}{2^m}, \frac{n+1}{2^m}) \cap [\frac{q}{2^m}, \frac{q+1}{2^m}) = \emptyset,$$

and, hence, $< \psi_{m,n}, \psi_{p,q} >= 0$.

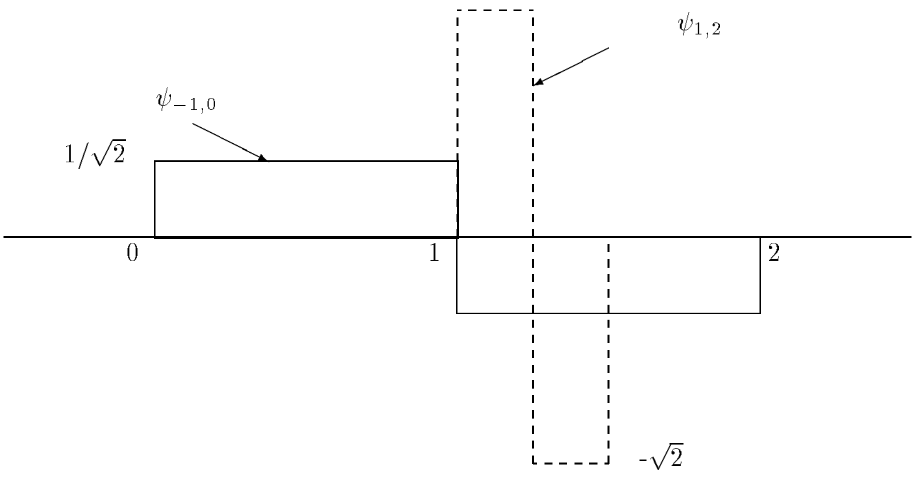
iii. Suppose $m \neq p$. Thinking of m, n, and p fixed, the intersection

$$[\frac{n}{2^m}, \frac{n+1}{2^m}) \cap [\frac{q}{2^p}, \frac{q+1}{2^p}) \tag{20}$$

is empty if q is large or small enough; and in such case we have $<\psi_{m,n}, \psi_{p,q}>= 0$.

If intersection (20) is non-empty, we use Proposition (5.3.2) to conclude that either $I_{m,n} \subseteq I_{p,q}$ or $I_{p,q} \subseteq I_{m,n}$. Suppose $I_{p,q} \subseteq I_{m,n}$. Then our proof of orthonormality would be complete if we could show that $\psi_{m,n}$ is constant when $\psi_{p,q}$ is nonzero. Before proving this generally in part v, we give the following example.

iv. Let $m = -1,\ n = 0,\ p = 1,\ q = 2$. Thus, $I_{-1,0} = [0, 2]$ and $I_{1,2} = [1, \frac{3}{2}]$. The relationship between $\psi_{-1,0}$ and $\psi_{1,2}$ is exhibited in Figure 0d



Note that $\psi_{-1,0}$ takes a constant value, viz, $-1/\sqrt{2}$, where $\psi_{1,2}$ is non-zero.

v. Without loss of generality, let $m < p$, assume the intersection in (20) contains an interval, and suppose

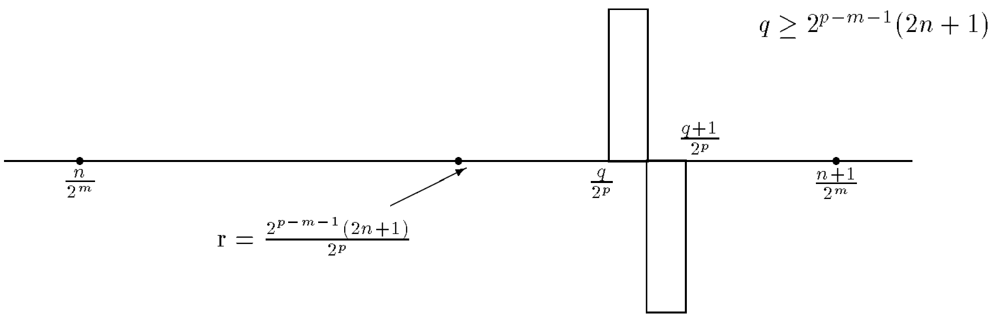
$$\frac{n}{2^m} < \frac{q}{2^p} < \frac{n+1}{2^m} \tag{21}$$

cf., case 3 in the proof of Proposition (5.3.2). From case 2 in the proof of Proposition (5.3.2) we know that $I_{p,q} \subseteq I_{m,n}$. Further, the midpoint r of $I_{m,n}$ is

$$r = \frac{1}{2}\left(\frac{n}{2^m} + \frac{n+1}{2^m}\right) = \frac{1}{2}\left(\frac{2n+1}{2^m}\right) = 2^{p-m-1}\left(\frac{2n+1}{2^p}\right).$$

Since $m < p$ then $m+1 \leq p$ and so $2^{p-m-1}(2n+1) \in \mathbb{Z}$. By (21), either $q < 2^{p-m-1}(2n+1)$ or $q \geq 2^{p-m-1}(2n+1)$.

In the first case, $q+1 \leq 2^{p-m-1}(2n+1)$ and so $\psi_{m,n} = 2^{m/2}$ when $\psi_{p,q}$ is non-zero. In the second case, $\psi_{m,n} = -2^{m/2}$ when $\psi_{p,q}$ is non-zero, e.g., Figure 5.3 □



Theorem 5.3.4 (The Haar Orthonormal Wavelet)
Let ψ be the Haar function. The Haar wavelet system $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.

Proof:

i. By the orthonormality of $\{\psi_{m,n}\}$, Proposition (5.3.3), and Theorem B [ON and span density $\Rightarrow$ ONB], it is sufficient to prove that $\overline{span}\{\psi_{m,n}\} = L^2(\mathbb{R})$.

In order to probe this density, let $\epsilon > 0$ and let $g \in L^2(\mathbb{R})$. We shall prove the following three facts, and these will complete the proof by the triangle inequality:

$$\exists N > 0 \quad \text{and} \quad \exists g_N \in L^2(\mathbb{R}) \qquad \text{such that } \operatorname{supp} g_N \subseteq [-2^N, 2^N] \tag{22}$$

and

$$g - g_N\|_{L^2(\mathbb{R})} < \frac{\epsilon}{3};$$

$$\exists M > 0 \quad \text{and} \quad \exists f \in L^2(\mathbb{R}) \qquad \text{such that } \operatorname{supp} f \subseteq [-2^N, 2^N], \tag{23}$$

$$f \quad \text{is constant on intervals of the form} \quad [\frac{n}{2^m}, \frac{n+1}{2^m}),$$

and

$$\|g_N - f\|_{L^2(\mathbb{R})} < \frac{\epsilon}{3};$$

$$\exists \{c_{m,n} \in \mathbb{C} : (m,n) \in F_1 \times F_2, F_j \subseteq \mathbb{Z} \quad \text{finite}\} \tag{24}$$

such that

$$\|f - \sum_{(m,n) \in F_1 \times F_2} c_{m,n} \psi_{m,n}\|_{L^2(\mathbb{R})} < \frac{\epsilon}{3}.$$

- Perhaps give (22) and (23) as exercises.
- Also, perhaps ask as an exercise the mechanicization (computer program) to prove (23) for a given g_N supported by $[-2^N, 2^N]$.

Should I match notation with MATLAB Wavelet Toolbox, and designate "e" (which follows) by "d" etc...?

ii. Let $f_0 = f$ and denote its constant value on $I_{M,n}$ as $f_0(\frac{n}{2^M}) \equiv f_0[n]$. We shall write f_0 as

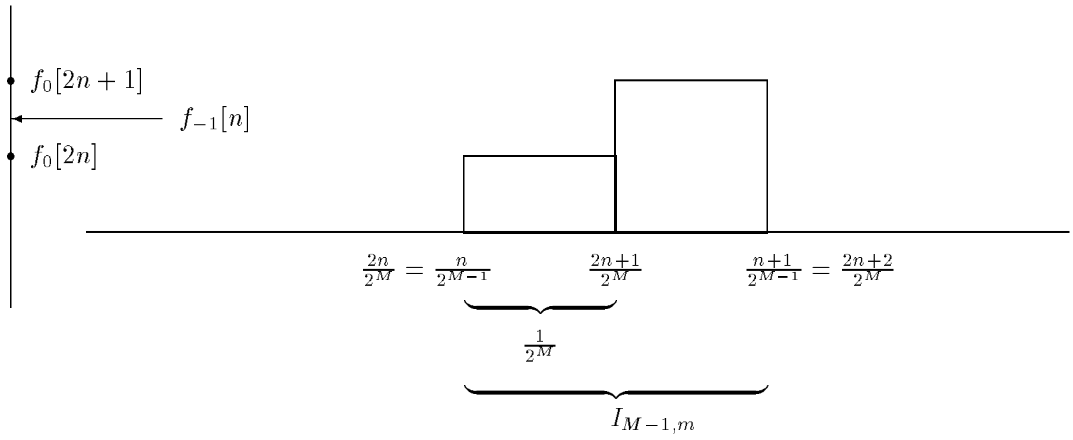
$$f_0 = f_{-1} + e_{-1}, \tag{25}$$

where f_{-1} will be an *approximation* of f_0 which will be piecewise constant on intervals of length 2^{-M+1}, i.e., twice the length of $I_{M,n}$.

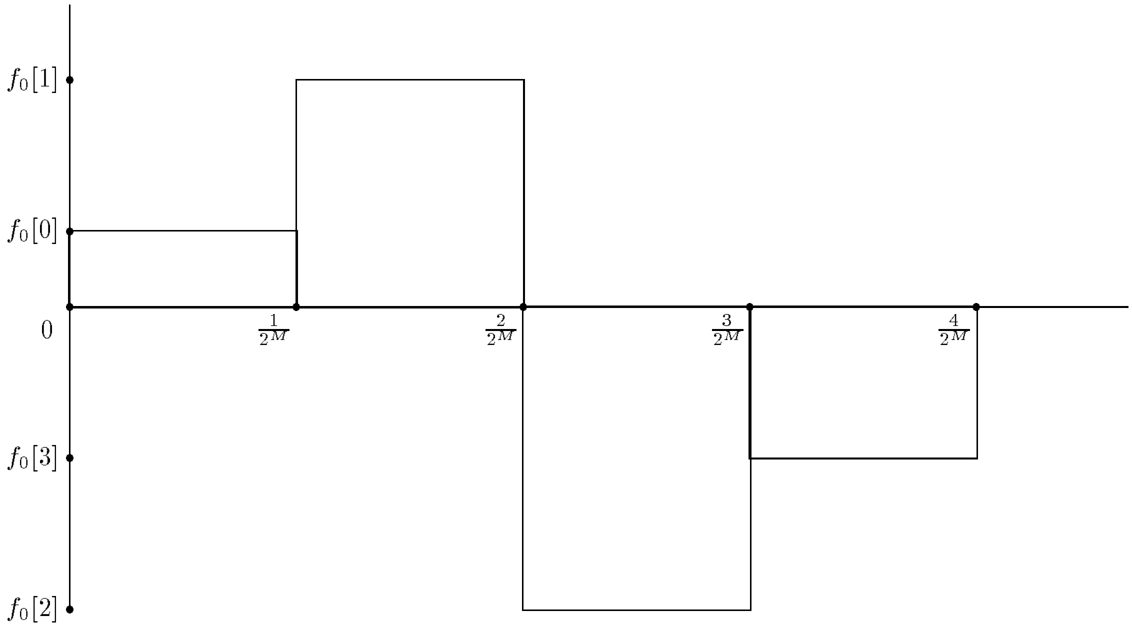
In order to define f_{-1}, define the constant value $f_{-1}[n]$ of f_{-1} on $I_{M-1,n}$ to be

$$f_{-1}[n] = \frac{1}{2}(f_0[2n] + f_0[2n+1]),$$

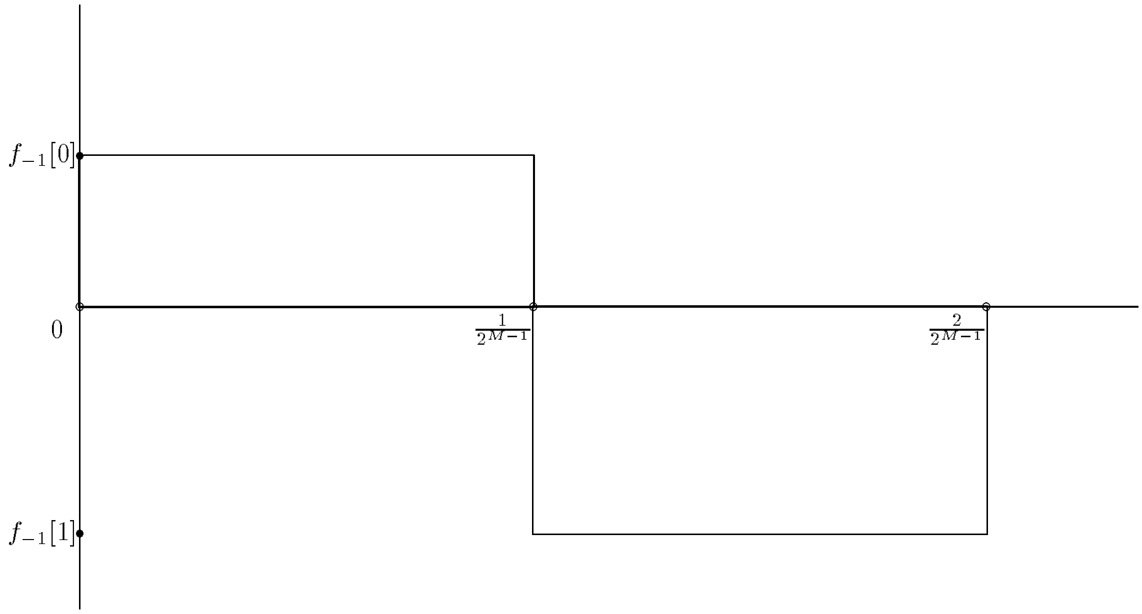
see Figure (0b).



Thus, if f_0 is defined as in Figure (0b),



then f_{-1} is defined as in figure (0b).



Now because of (25) ($e_{-1} = f_0 - f_{-1}$), we see that e_{-1} has the same step-size as f_0, viz., $1/2^M$. The value of e_{-1} on $I_{M,n}$ is $e_{-1}(n/2^M)$ and it will be designated $e_{-1}[n]$.
If n is even, i.e., $n = 2k$, then

(26)
$$\begin{aligned} e_{-1}[2k] &= e_{-1}\left(\tfrac{2k}{2^M}\right) = f_0\left(\tfrac{2k}{2^M}\right) - \tfrac{1}{2}\left(f_0\left(\tfrac{2k}{2^M}\right) + f_0\left(\tfrac{2k+1}{2^M}\right)\right) \\ &= f_0[2k] - \tfrac{1}{2}(f_0[2k] + f_0[2k+1]) \\ &= \tfrac{1}{2}(f_0[2k] - f_0[2k+1]). \end{aligned}$$

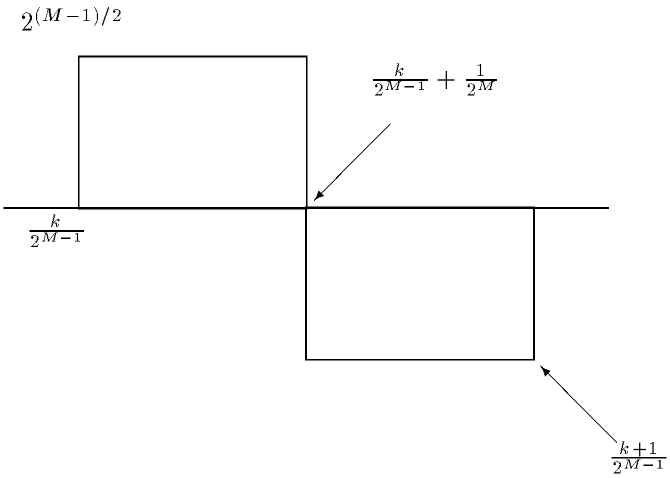
Similarly, if n is odd, i.e., $n = 2k+1$, then

(27)
$$\begin{aligned} e_{-1}[2k+1] &= e_{-1}\left(\tfrac{2k+1}{2^M}\right) = f_0\left(\tfrac{2k+1}{2^M}\right) - \tfrac{1}{2}\left(f_0\left(\tfrac{2k}{2^M}\right) + f_0\left(\tfrac{2k+1}{2^M}\right)\right) \\ &= f_0[2k+1] - \tfrac{1}{2}(f_0[2k] + f_0[2k+1]) \\ &= \tfrac{1}{2}(f_0[2k+1] - f_0[2k]) = -e_{-1}[2k]. \end{aligned}$$

Recall that $\psi_{m,n}(t) = 2^{m/2}\psi(2^m t - n)$, i.e.,

$$\psi_{m,n}(t) = \begin{cases} 2^{m/2} & \text{if} \quad 2^m t - n \in [0, \tfrac{1}{2}), \quad \text{i.e.,}\ 2^m t \in [n, n+\tfrac{1}{2}), \quad \text{i.e.,}\ t \in \left[\tfrac{n}{2^m}, \tfrac{n}{2^m} + \tfrac{1}{2^{m+1}}\right), \\ -2^{m/2} & \text{if} \quad 2^m t - n \in [\tfrac{1}{2}, 1), \quad \text{i.e.,}\ 2^m t \in [n+\tfrac{1}{2}, n+1), \quad \text{i.e.,}\ t \in \left[\tfrac{n}{2^m} + \tfrac{1}{2^{m+1}}, \tfrac{n+1}{2^m}\right), \\ 0 & \text{otherwise.} \end{cases}$$

Thus $\psi_{M-1,k}$ appears as in Figure 0b.



From (27), we obtain

$$e_{-1}[2k]\psi_{M-1,k}(t) = \begin{cases} 2^{(M-1)/2}e_{-1}[2k] & \text{on} \quad \left[\frac{k}{2^{M-1}}, \frac{k}{2^{M-1}} + \frac{1}{2^M}\right), \\ 2^{(M-1)/2}e_{-1}[2k+1] & \text{on} \quad \left[\frac{k}{2^{M-1}} + \frac{1}{2^M}, \frac{k+1}{2^{M-1}}\right). \end{cases}$$

Hence,

$$e_{-1} = \sum_{k=-2^{N+M-1}}^{2^{N+M-1}-1} 2^{-(M-1)/2} e_{-1}[2k]\psi_{M-1,k} \quad \text{on} \quad [-2^N, 2^N]. \tag{28}$$

To see that these are the proper limits of summation, let $k = -2^{N+M-1}$ and so the left endpoint of the support of $\psi_{M-1,k}$ is (Figure 0b)

$$\frac{-2^{N+M-1}}{2^{M-1}} = -2^N.$$

Similarly, for $k = 2^{N+M-1} - 1$, the right endpoint of the support of $\psi_{M-1,k}$ is (Figure 0b)

$$\frac{2^{N+M-1} - 1}{2^{M-1}} + 1 = 2^N.$$

From (25) and (28), we have

$$f = f_0 = f_{-1} + \sum_k c_{M-1,k}\psi_{M-1,k} \tag{29}$$

where f_{-1} is of the same type as f_0 but with stepwidth twice as long. The finite sum over k in (28), is e_{-1} and $\operatorname{supp} e_{-1} = [-2^N, 2^N]$.

iii. We apply the same procedure to f_{-1} as we did to f_0 and we form f_{-2} and e_{-2}. Consequently, we obtain

$$f_{-1} = f_{-2} + \sum_k c_{M-2,k}\psi_{M-2,k},$$

where f_{-2} is piecewise constant on intervals of length twice that of those for f_{-1}. We continue in this way forming f_{-3}, f_{-4},... Note that the origin plays a role in this algorithm, i.e., we are "doubling from the origin" going both left and right, e.g., Figures 0b and 0b.

We see that after $M+N$ steps we have

$$f = f_0 = f_{-(M+N)} + \sum_{m=1}^{M+N} \sum_k c_{M-m,k}\psi_{M-m,k} \tag{30}$$

In fact, from step 1, we have

$$f_0 = f_{-1} + \sum_k c_{M-1,k}\psi_{M-1,k}$$

and in step 2, we have

$$f_{-1} = f_{-2} + \sum_k c_{M-2,k}\psi_{M-2,k};$$

and, hence,

$$f_0 = f_{-2} + \sum_k c_{M-1,k}\psi_{M-1,k} + \sum_k c_{M-2,k}\psi_{M-2,k}.$$

In step 3,

$$f_{-2} = f_{-3} + \sum_k c_{M-3,k}\psi_{M-3,k};$$

and, hence,

$$f_0 = f_{-3} + \sum_{m=1}^{3} \sum_k c_{M-m,k}\psi_{M-m,k}$$

In this way we obtain (30).

To see that the sum in (30) goes from $m = 1$ to $M + N$, notice we have proceeded in the following way.

At step 0, $f = f_0$ is constant on $2 \times 2^{M+N} = 2^{M+N+1}$ intervals. On the non-negative axis these intervals are

$$\left[0, \frac{1}{2^M}\right), \left[\frac{1}{2^M}, \frac{2}{2^M}\right), \ldots, \left[\frac{2^{M+N}-1}{2^M}, \frac{2^{M+N}}{2^M}\right),$$

and there are 2^{M+N} of them.

At step 1, f_{-1} is constant on $2 \times 2^{-1} \times 2^{M+N} = 2^{M+N}$ intervals. On the non-negative axis these intervals are

$$\left[0, \frac{1}{2^{M-1}}\right), \left[\frac{1}{2^{M-1}}, \frac{2}{2^{M-1}}\right), \left[\frac{2}{2^{M-1}}, \frac{3}{2^{M-1}}\right), \ldots, \left[\frac{2^{M+N-1}-1}{2^{M-1}}, \frac{2^{M+N-1}}{2^{M-1}}\right),$$

i.e.,

$$\left[0, \frac{2}{2^M}\right), \left[\frac{2}{2^M}, \frac{4}{2^M}\right), \ldots, \left[\frac{2^{M+N}-2}{2^M}, \frac{2^{M+N}}{2^M}\right),$$

and there are 2^{M+N-1} of them.

At step 2, f_{-2} is constant on $2 \times 2^{-2} \times 2^{M+N} = 2^{M+N-1}$ intervals. On the non-negative axis these intervals are

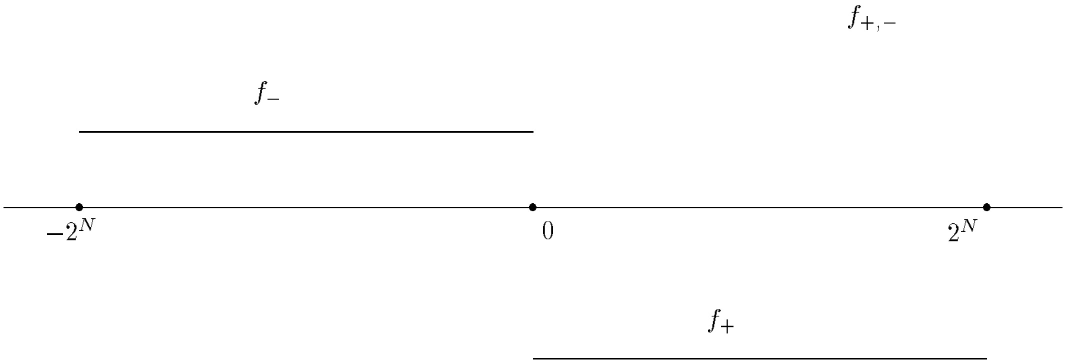
$$\left[0, \frac{1}{2^{M-2}}\right), \left[\frac{1}{2^{M-2}}, \frac{2}{2^{M-2}}\right), \ldots, \left[\frac{2^{M+N-2}-1}{2^{M-2}}, \frac{2^{M+N-2}}{2^{M-2}}\right),$$

i.e.,

$$\left[0, \frac{4}{2^M}\right), \left[\frac{4}{2^M}, \frac{8}{2^M}\right), \left[\frac{8}{2^M}, \frac{12}{2^M}\right), \ldots, \left[\frac{2^{M+N}-4}{2^M}, \frac{2^{M+N}}{2^M}\right).$$

Proceeding in this way, we see that at step $M+N$, $f_{-(M+N)}$ is constant on $2^{(M+N)-(M+N)}-1+1 = 1$ interval on the non-negative axis, viz., $[0, 2^N)$. Similarly, $f_{-(M+N)}$ is constant on $[-2^N, 0)$.

Notationally, let $f_{+,-} = f_{-(M+N)}$ and write f_+, resp., f_-, for $f_{+,-}$ on $[0, 2^N)$, resp., $[-2^N, 0)$, see Figure 0c



Recalling the definition of f_{-1} and proceeding recursively, we see that f_+ and f_- are averages over $[0, 2^N)$ and $[-2^N, 0)$, resp. (Exercise 5.)

iv. Now suppose we continue this averaging procedure. For example, on the non-negative axis, consider the "double" interval $[0, 2^{N+1})$. In this case, the average of f on $[0, 2^{N+1})$ is the constant $2^{-1}f_+$. In the process of computing this average, we must also introduce the Haar functions

$$\psi_{-(N+1),0} \quad \text{and} \quad \psi_{-(N+1),-2^{(N+1)}},$$

and hence are able to obtain the decomposition

$$f = f_0 = 2^{-1} f_{+,-} + \sum_{m=1}^{M+N+1} \sum_k c_{M-m,k}\psi_{M-m,k} \quad \text{on} \quad [-2^{N+1}, 2^{N+1}).$$

Continuing in this way, we obtain

$$f = f_0 = 2^{-P} f_{+,-} + \sum_{m=1}^{M+N+P} \sum_k c_{M-m,k}\psi_{M-m,k} \quad \text{on} \quad [-2^{N+P}, 2^{N+P}). \tag{31}$$

In fact, (31) is true on all of $\mathbb{R}$ since all of the functions in (31) vanish outside of $[-2^{N+P}, 2^{N+P})$.

v. We compute

$$\|2^{-P} f_{+,-}\|^2_{L^2(\mathbb{R})} \leq 2^{-2P} \int_{-2^{N+P}}^{2^{N+P}} \|f_{+,-}\|^2_{L^\infty(\mathbb{R})} dt \tag{32}$$

$$= 2^{N+P+1-2P}||f_{+,-}||^2_{L^\infty(\mathbb{R})} = c2^{(N+1)-P}.$$

Choose $P > 0$ such that $c2^{(N+1)-P} < \left(\frac{\epsilon}{3}\right)^2$. Then (31) and (32) combine to give

$$||f - \sum_{m=1}^{M+N+P} \sum_k c_{M-m,k}\psi_{M-m,k}||_{L^2(\mathbb{R})} < \frac{\epsilon}{3}, \tag{33}$$

noting that the sums over k are finite sums, e.g., (28). The inequality (33) completes the proof of (24). □

• Give example to show how the proof of Theorem (5.3.4) is a primordial case of an MRA. Make this the subject of Example (5.3.5).
• The construction in Theorem (5.3.4) can be programmed and is precisely the DWT Mallat algorithm. In fact, at each step we obtain averages and errors (details) for a given average, the first "average" at step 0 being the given function f. Then we reconstruct f in terms of an average at step K and all the detail from steps $k = 1, ..., K$.
• Note that $\int \sum c_{m,n}\psi_{m,n} = 0$ even though the original function f approximated by $\sum c_{m,n}\psi_{m,n}$ does not necessarily have the property that $\int f(t)dt = 0$.

Example 5.3.5 See Walker in [[]NAMS vol44 no.6 pp658 ff] for his remarks on Haar (at first glance it seemed like nothing new.)

Example 5.3.6 Note that the approximation 24 in the proof of Theorem 5.3.4 is not true for $L^1(\mathbb{R})$ since the L^1 estimate corresponding (32) and 33 fails. In fact,

$$||2^{-P}f_{+,-}||_{L^1[0,\infty)} = 2^{-P}\int_0^{2^{N+P}} |f_+|dt = |f_+|2^N.$$

Should I put L^p, $p > 0$, material here? These notes are a part of a manuscript I am working on, as well as a PhD project. Reference this work and role in ATDR and Image processing. Should I cross reference with section 1.8? Put my $p > 0$ notes in [[]SUP(4)].

Theorem 5.3.7 Let ψ be the Haar function and define

$$\psi_{m,an}(t) = 2^{m/2}\psi(2^m t - an)$$

for $a > 0$ and $m, n \in \mathbb{Z}$

a. If a __________ then $\{\psi_{m,an} : m, n \in \mathbb{Z}\}$ is a frame with bounds $A =$ ____, $B =$ ____.

b. If a __________ then $\{\psi_{m,an} :\ m, n \in \mathbb{Z}\}$ is a tight frame with bounds $A = B =$ ____.

c. If a __________ then $\{\psi_{m,an} :\ m, n \in \mathbb{Z}\}$ is not a frame.

d. $\{\psi_{m,an} :\ m, n \in \mathbb{Z}\}$ is an exact frame if and only if ________________.

<u>Check problem from Fall 1995 math 648.</u> Insert exercise on this.

Perhaps give a difficult exercise. Was it completed by Chui and Shi?

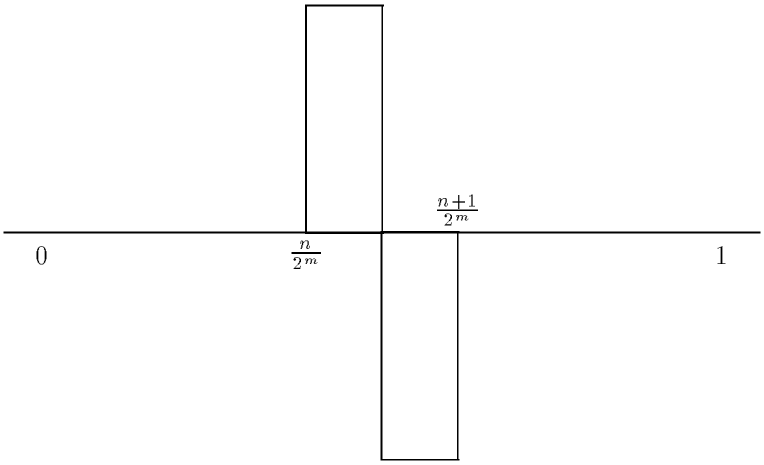
See Chui and Shi : Proc. AMS ($a = \frac{1}{N}, N$ odd) c. 1994? and ACHA c. 1994 and Advances in Computational Math, Dik... and C.A.Micchelli, eds, pp139-155, 1994 World Scientific, (for d-dim case). See what SHIDONG did in his thesis on this, and quote it.

Example 5.3.8 (The Haar ONB for $L^2(\mathbb{T})$)

a. Let $\psi = 1$ on $[0, 1)$ and let ψ be the Haar function on $\mathbb{R}$. It turns out that

$$\{\phi, \psi_{m,n} : m = 0, 1, \ldots \quad \text{and} \quad 0 \le n < 2^m\} \tag{34}$$

is an ONB *for* $L^2[0, 1)$ (Exercise 5.33). Note that if $m \in 0, 1, ..$ and $0 \le n < 2^m$, then $\psi_{m,n}$ appears as in Figure 1. In fact, $0 \le n/2^m < (n + 1)/2^m \le 1$.



Thus, by extending the functions in (34) 1-periodically on $\mathbb{R}$, we can assert that the *sequence defined by 34*
is an ONB for $L^2(\mathbb{T})$. This result is due to Haar.

Give page number in Haar's (1908) paper/ thesis.

Draw $\psi_{m,n}$ for specific $m = 0, 1, \ldots$ and specific $0 \leq n < 2^m$.

b. i. The Haar system on $[0, 1)$ is a basis for $L^1[0, 1)$, e.g., [[]Marti pp49-50], [[]Shipp, Wade, and Simon, p. 203].

ii. There are no conconditional bases (define it, it's in section 5.1 notes) for $L^1[0, 1)$ [Singer 1970, section 15 part II]

iii. Is $\{\tau_n bf 1_{[0,1)}\} \cup \{\psi_{m,n}\}$, where ψ is the Haar function, a basis for $L^1(\mathbb{R})$?

Check my research notes on bases and projections from 1996 for more updated historical data on this subject.

c. If $f \in L^p(\mathbb{T})$, $1 < p < \infty$, and its sequence of wavelet coefficients is $\{< f, \psi_{m,n} >: m = 0, 1, \ldots$ and $0 \leq n < 2^m\}$, then if $g \in L^1(\mathbb{T})$ has the property that

$$\forall m = 0, 1, \ldots \quad \text{and} \quad 0 \leq n < 2^m, | < f, \psi_{m,n} > | = | < g, \psi_{m,n} > |,$$

we can conclude that $g \in L^p(\mathbb{T})$.

Give precise reference from Yves Meyer.

Such a result is only true for Fourier coefficients in the case $p = 2$.

Should I add more about the subtlety of Fourier series, or refer to the deLeeuw, Kahane, Katznelson theorem quoted on p. 174?

Note the importance of positive numbers in applications. Refer back to interferometer, intensity readings in CAT scanners, spectrograms, etc. The above result is in the spirit of deducing precise L^p-average information about a signal g if we know amplitude values.

d. In section 5.2.13, we saw how to generate multidimensional wavelet ONB for $L^2(\mathbb{R}^d)$. In the case of the Haar function ψ and the characteristic function $\phi = \mathbf{1}_{[0,1)}$, we have the following situation for $L^2(\mathbb{T}^d)$.

Example 5.3.9 *Rectilinear tensor product wavelet decomposition.*

Since

$$\phi(2^{j_1}x_1 - k_1)\phi(2^{j_2}x_2 - k_2)\ldots\phi(2^{j_d}x_d - k_d)$$

is the characteristic function of

$$\frac{k_1}{2^{j_1}} \le x_1 \le \frac{k_1+1}{2^{j_1}}, ..., \frac{k_d}{2^{j_d}} \le x_d \le \frac{k_d+1}{2^{j_d}} \tag{35}$$

these functions span $L^2(\mathbb{R}^d)$.
Let $\mathbf{j} = (j_1, j_2, ..., j_d), \mathbf{k} = (k_1, k_2, ..., k_d)$, and consider the set of functions

$$\Psi_{\mathbf{j},\mathbf{k}}(\mathbf{x}) = \psi_{j_1,k_1}(x_1)\psi_{j_2,k_2}(x_2)...\psi_{j_d,k_d}(x_d), \tag{36}$$

where $\mathbf{j} \in \mathbb{Z}^d$, $\mathbf{k} \in \mathbb{Z}^d$.
Also, we showed that the functions in (36) form an ONB for $L^2(\mathbb{R}^d)$. If $\{\mathbf{j_i}\}$ runs over only the non-negative integers, and $0 \le k_i < 2^{j_i}$, then functions in (36) form an orthonormal basis for a subspace of $L^2(\mathbb{T}^d)$. In order to obtain an ONB for $L^2(\mathbb{T}^d)$, the function in (36) must be supplemented by functions of the form

$$\phi_{x_{i_1}}...\phi_{x_{i_l}}\psi_{j_{i_{l+1}},k_{i_{l+1}}}(x_{i_{l+1}})...psi_{j_{i_d},k_{i_d}}(x_{i_d})$$

for $l = 1, 2, ..., d$.
Multiresolution analysis basis As before $\phi(2^j x_1 - k_1)\phi(2^j x_2 - k_2)...\phi(2^j x_d - k_d)$ is the characteristic function of

$$\frac{k_1}{2^j} \le x_1 \le \frac{k_1+1}{2^j}, ..., \frac{k_d}{2^j} \le x_d \le \frac{k_d+1}{2^j},$$

and these functions form a multiresolution analysis $\{V_j\}$ of $L^2(\mathbb{R}^d)$.
Further, as mentioned in Section 5, and as in the multidimensional case, we can write

$$V_{n+1} = W_n \oplus V_n,$$

but now W_n has $2^d - 1$ different wavelet functions and their translates as its basis. To obtain these functions, we consider each non-empty subset F of $\{0, 1, ..., d\}$, and associate the function

$$\prod_{i\in F}(2^{j/2}\psi(2^j x_i - k_i)) \times \prod_{i\in F^c}(2^{j/2}\phi(2^j x_i - k_i)). \tag{37}$$

The empty set is not used, since then all the univariate functions in the product would be ϕ functions, and so the resulting function on $\mathbb{R}^d$ would be in V_j.
This basis can be extended to $L^2(\mathbb{T}^d)$, where again we only need consider non-negative j's and $0 \le k_i \le 2^j - 1$. To complete the bases we need to include $\phi(x_1)\phi(x_2)...\phi(x_d)$ along with those of the form (37).

- **Rewrite part *d*. Drop the bold-face? (used in Ch. 2). I don't use bold in discussing $n \in \mathbb{Z}^d$ in Section 5.5. Perhaps drop bracketed material ahead.**
- F^c **ok?**
- **Don't use "extended" in last sentence.**
- **Do details for this more carefully.**
- **Compare with multidimensional calcs in Section 5.2.**

Example 5.3.10 (Haar Series)

> **Quote from J. R. McLaughlin, Haar series, Transactions of the American Mathematical Society, 137 (1969), pp 153-176 on "Haar-Fourier Series". Very good bibliography.**

> **See McLaughlin Pacific Journal of Math 27 (2968) for Rademacher sums. Also see Keogh Journal of the London Mathematical Society 33 (1958) 454-455. I have Keogh in Littlewood volume.**

Let $\theta_p : \mathbb{T} \to \mathbb{C}$ be the Haar system defined on $[0,1)$ as follows:

$$\theta_0 = 1 \text{ on } [0,1)$$

and, for $p = 2^m + n$ with $0 \leq n < 2^m, m = 0, 1, \ldots,$

$$\theta_p(t) = \begin{cases} 2^{m/2} & \text{if } t \in [\frac{n}{2^m}, \frac{n+\frac{1}{2}}{2^m}) \\ -2^{m/2} & \text{if } t \in [\frac{n+\frac{1}{2}}{2^m}, \frac{n+1}{2^m}) \\ 0 & \text{if } t \in [0,1) \setminus [\frac{n}{2^m}, \frac{n+1}{2^m}). \end{cases}$$

This defines the same set of functions as in Example 5.4.8a, cf., Example A.11.

Analogous to Fourier series, we can formally define the *Haar series* of $F \in L^1(\mathbb{T})$ as the series

$$H(F)(\gamma) = \sum_{p=0}^{\infty} f[p]\theta_p(\gamma),$$

where

$$\forall p = 0, 1, \ldots, \quad f[p] = \int_{\mathbb{T}} F(\gamma)\theta_p(\gamma)d\gamma.$$

Haar [Haa09, pages xxx] proved the following results for $F \in L^1(\mathbb{T})$:

i. $H(F) = F$ a.e. ;

ii. if F is continuous at γ, then $H(F)(\gamma) = F(\gamma)$;

iii. if F is uniformly continuous on an interval $I \subseteq \mathbb{T}$, then $H(F)$ converges uniformly to F on I (obviously, continuity is a sufficient hypothesis for closed intervals).

As such, the subtleties associated with Carleson's Theorem for Fourier series do not arise for Haar series, cf., Section 3.2.8.

On the other hand, the unboundedness of the Haar system leads to other problems. For example, if $F \in C(\mathbb{T})$ and $\sum_{p=0}^{\infty} |f[p]| < \infty$, can one assert that $\sum_{p=0}^{\infty} |f[p]\theta_p(\gamma)|$ converges uniformly on $\mathbb{T}$? It was shown by

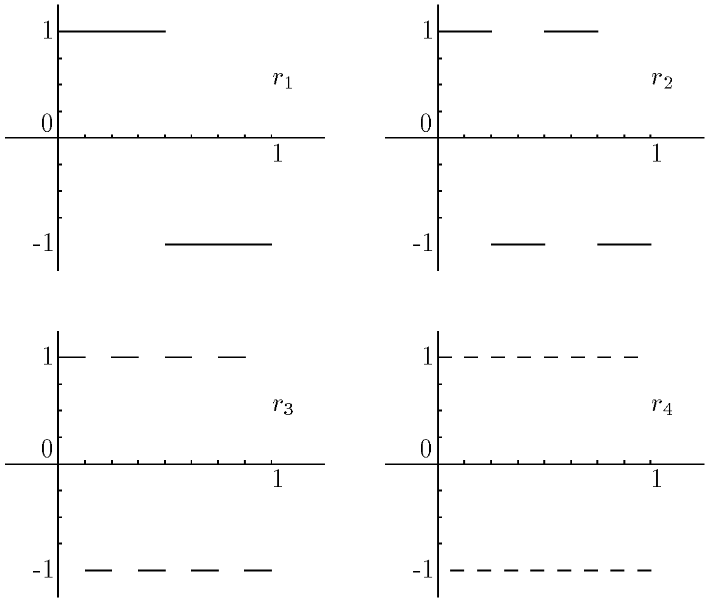


Figure 25: Rademacher Functions r_1, r_2, r_3, r_4

McLaughlin [McL69] that the answer is "no". There have been a great many contributors to the theory of Haar series, and pointwise results. These include Ciesielski and Musielak, Faber, Golubov, Matveev, and Ul'janov, e.g., [Ulj64].

Definition 5.3.11 (Walsh Functions)

a. The sequence $\{r_n : n \in \mathbb{N}\}$ of *Rademacher functions* on $\mathbb{T}$ is defined by the 1-periodic functions

$$\forall n = 1, 2, \ldots,\ r_n(t) = \operatorname{sgn}(\sin 2^n \pi t)$$

on $\mathbb{R}$, see *Figure 1* for $r_1, \ldots, r_4$ on $[0, 1)$. The n^{th} Rademacher function has 2^{n-1} cycles each of length $\frac{1}{2^{n-1}}$ on $[0, 1)$.

b. It is easy to see that $\{r_n\}$ is orthonormal. On the other hand, $\{r_n\}$ is not an ONB for $L^2(\mathbb{T})$. In fact, $\langle \cos 2\pi t, r_n(t) \rangle \equiv \int_{\mathbb{T}} \cos 2\pi t\ r_n(t)\ dt = 0$ for each n, and this proves that $\overline{\operatorname{span}}\{r_n\} \neq L^2(\mathbb{T})$ by the Hahn-Banach Theorem.

c. The sequence $\{w_n : n = 0, 1, \ldots\}$ of *Walsh functions* on $\mathbb{T}$ is defined by $w_0 = 1$ and the other 1-periodic functions

$$\forall n = 1, 2, \ldots,\ w_n(t) = \prod_{\epsilon_j = 1} r_j(t),$$

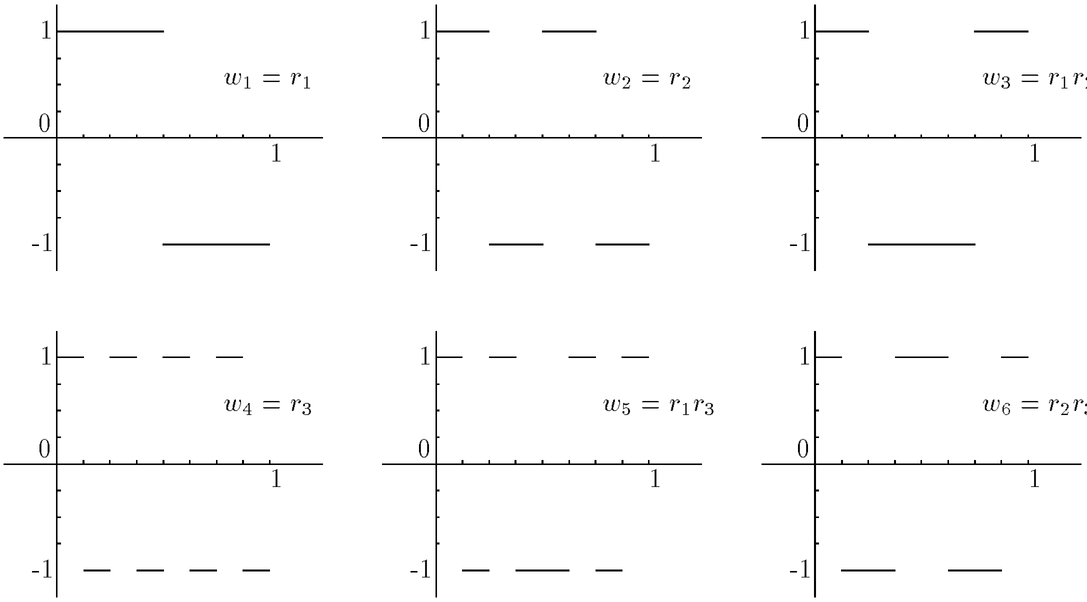


Figure 26: Walsh Functions $w_1, w_2, \ldots, w_6$

on $\mathbb{R}$, where $n = \sum_{j=1}^{\infty} \epsilon_j 2^{j-1}$ is the binary expansion of n, cf., Remark 3.9.4 d. *Figure 3* exhibits $w_1, \ldots, w_6$.

Remark 3.9.4 d will be renumbered in HAA(2nd).

$5 = \epsilon_1 \cdot 2^0 + \epsilon_2 \cdot 2^1 + \epsilon_3 \cdot 2^2 + \epsilon_4 \cdot 2^3 + \ldots$. **Thus** $\epsilon_1 = \epsilon_3 = 1$ **and** $\epsilon_j = 0$ **otherwise. Thus** $w_5 = r_1 \cdot r_3$**.**

Clearly, each Rademacher function is a Walsh function. In fact, $W_{2^{k-1}} = r_k, k \in \mathbb{N}$.

d. Walsh function have had many applications in classical signal processing, e.g. [Harm69], [SS74].

Should I add my notes from my work on Walsh functions at MITRE? (See my notes on sequency.) Go through above books again for good applications. Recall that Harmuth introduced sequency in 1968 and it can be viewed as a generalization of frequency, e.g., [SS74, page 323].

Note there is a large bibliography on Walsh functions in [SS74, pages 417-460].

I have the 1970 and 1971 Proceedings of Applications of Walsh Function sponsored by NRL and the University of Maryland. Joe Walsh gave me that 1971 proceedings in 1972, and discussed them with me; but at the time I was doing other things, and I didn't pursue it until I went to MITRE.

Walsh has written [1970 Proceedings on Applications of Walsh Function] that Rademacher and Walsh functions were discovered independently. Rademacher's manuscript was received by the editors of Mathematische Annalen on October 8, 1921. Walsh presented his functions to the American Mathematical Society at a meeting on February 25, 1922, and published an abstract of his paper in Bulletin AMS, 28 (1922), 241.

In this discourse, Walsh also noted Franklin's paper [Fra28], where Franklin obtained orthonormal functions by orthogonalizing the integrals of the Haar function, thereby obtaining a sequence of continuous orthonormal functions in terms of which an arbitrary continuous function can be uniformly expanded. It is this Franklin system along with higher order spline systems, which have been used as unconditional bases for H^p spaces, and which was the starting point for Jan-Olov Strömberg's celebrated "prewavelet" paper [Str83], e.g., [Mey90].

I had some ideas on the Franklin system, which I marked on my copy of his paper. Should I quote Fine's article as well as Hirschman's AMS memoir. I have Hirschman (see page 28).

See the discussion of channel crosstalk and the work by engineers on Walsh functions, around the turn of the century, i.e., 1900. This is in one of the Walsh-NRL-UMD books I have.

Do Rudin-Shapiro polynomials vis a vis Walsh functions. See my notes in draft of my manuscript with Erica Bernstein. See George Benke's article in JFAA, volume 1. See J. Byrnes and/or H. Shapiro manuscripts, especially with regard to antennas and/or uncertainty. Also note value of ± 1 sequences in information theory, spread spectra, etc. See original 40s (or 50s?) paper on Rudin-Shapiro polynomials.

Have I already mentioned that Franklin was Wiener's brother-in-law?

Theorem 5.3.12 (Walsh Theorem (1923))

The sequence $\{w_n : n = 0, 1, \ldots\}$ of Walsh functions is an ONB for $L^2(\mathbb{T})$.

The proof is not difficult. First show orthonormality and then prove ONB by showing $\overline{\text{span}}\{w_n\} = L^2(\mathbb{T})$. To this end, let $f \in L^2[0,1)$, define

$$g(t) = \int_0^t f(u)du,$$

and suppose $\int_0^1 f(u)w_n(u)du = 0$ for all $n \geq 0$. Use this hypothesis to prove that $g = 0$ on the set of dyadic rationals. Since these are dense in $[0,1)$, we have that $g = 0$ on $[0,1)$. Consequently, $f = 0$ a.e.

| **This is Exercise 5.18.**

$$0 = \int_0^1 f(u)w_1(u)du = \int_0^{\frac{1}{2}} f(u)du - \int_{\frac{1}{2}}^1 f(u)du = g(\frac{1}{2}) - (g(1) - g(\frac{1}{2}))$$
$$= 2g(\frac{1}{2}) - g(1) = 2g(\frac{1}{2}) - \int_0^1 f(u)w_0(u)du = 2g(\frac{1}{2}).$$

Thus, $g(\frac{1}{2}) = 0$.

Remark 5.3.13 (Walsh Functions and the Walsh Group)

| **See [Edw67, Chapter 14], [SWS90]**

a. A well known theorem, going back to Gauss, asserts that *any commutative finite group G is isomorphic to a direct product,*

$$\mathbb{Z}_{N_1} \times \cdots \times \mathbb{Z}_{N_n},$$

where each N_j is a power of a prime number, e.g., [???].

Let G be a locally compact abelian group (LCAG), i.e., a commutative group endowed with a Hausdorff locally compact topology under which the mapping

$$G \times G \longrightarrow G$$
$$(x, y) \longmapsto x - y$$

is continuous. (We have written the group operation additively.)

A *character* of a LCAG G is a continuous homomorphism

$$\gamma : G \to \{z \in \mathbb{C} : |z| = 1\},$$

the range being the multiplicative group of unimodular complex numbers taken with the topology (convergence criterion) from $\mathbb{C}$, cf., Definition 3.8.1 d.

| **Section 3.8 will be in Chapter 4 of HAA2nd.**

Thus writing $(x, \gamma) \equiv \gamma(x)$ for $x \in G$ and a character γ, we have

(38) $$\forall x, x_1, x_2 \in G,\ |(x, \gamma)| = 1 \text{ and } (x_1 + x_2, \gamma) = (x_1, \gamma)(x_2, \gamma).$$

$\widehat{G}$ denotes the set of characters of G, and $\widehat{G}$ itself becomes a LCAG, where addition is defined by

(39) $$\forall x \in G \text{ and } \forall \gamma_1, \gamma_2 \in \widehat{G}, (x, \gamma_1 + \gamma_2) = (x, \gamma_1)(x, \gamma_2).$$

The locally compact topology on $\widehat{G}$, as well as many other properties of LCAGs, are presented in [Rud62]. Generally,

$$\widehat{\widehat{G}} = G. \tag{40}$$

It is well-known that $\widehat{\mathbb{R}} = \mathbb{R}$, $\widehat{\mathbb{Z}} = \mathbb{T}$, and $\widehat{\mathbb{T}} = \mathbb{Z}$. It is also not difficult to prove that *if G is a commutative finite group then $\widehat{G} = G$.*

Refer to exercise where I relate "integral" transforms to topological algebras, viz., Exercise 2.41

b. Now consider the finite group

$$W_n = \mathbb{Z}_2 \times \cdots \times \mathbb{Z}_2,$$

with n factors in the direct product. "W" stands for Joseph L. Walsh (1895-1973), of Harvard University and the University of Maryland. We define the mapping

$$W_n \longrightarrow \{0, 1, 2, 3, \ldots, 2^n - 1\} \tag{41}$$

$$x = (\epsilon_1, \ldots, \epsilon_n) \longmapsto \sum_{j=1}^{n} \epsilon_j 2^{j-1}, \tag{42}$$

where each $\epsilon_j \in \{0, 1\}$. The mapping is bijective, and we transfer the group operation on W_n to $\{0, 1, \ldots, 2^n - 1\}$. As such, $\{0, 1, \ldots, 2^n - 1\}$ can be considered as a group isomorphic to W_n.

Be more explicit in writing down that W_n is isomorphic to $\mathbb{Z}_{2^n}$.

In part c we use the term *complete direct sum* to mean the group defined as an infinite product of groups in which addition is performed coordinatewise, cf., our discussion of direct sums in Sections 5.? and 5.?

J. L. Walsh was mentioned earlier. Consolidate his history.

c. For our next step we extend W_n, and we define the complete direct sum

$$W = \mathbb{Z}_2 \times \mathbb{Z}_2 \times \cdots .$$

Since each component $\mathbb{Z}_2$ is a commutative group, then W is a commutative group. As before, we assume $\mathbb{Z}_2$ is endowed with the discrete topology, and it is a compact group. We then endow W with the product topology. (The *product topology* is generated by open sets of the form

$$U = U_1 \times U_2 \times \cdots \subseteq W,$$

where $U_j = \mathbb{Z}_2$ for all but finitely many indices, and none of the U_j is empty.) By the Tychonov Theorem, W *is a compact abelian group*, e.g., [Ben76], [Rud62], [Rud73], [Tay58]. W is called the *Walsh group* or the *dyadic group.*

Explicitly express convergence in this topology.

Analogous to the case of W_n, we define the correspondence

$$(43)\qquad \begin{aligned} J : W &\longrightarrow [0,1] \\ x = (\epsilon_1, \ldots) &\longmapsto \sum_{j=1}^{\infty} \epsilon_j 2^{-j}, \end{aligned}$$

where each $\epsilon_j \in \{0,1\}$. J is not injective, since, for example,

$$J((1,0,0,0,\ldots)) = \frac{1}{2} = J((0,1,1,1,\ldots)).$$

This "duplicity" is not a serious problem, and a modified version of J, which we also denote by J, is a bijectiion $W \to [0,1)$, see Exercise 5.40.

As an Exercise, ask to compare the topology of W and the topology of $[0,1)$.
Also compare the addition in W with addition mod 1 in $[0,1)$
Instead of the mapping (43), consider the mapping

$$\begin{aligned} J_3 : W &\longrightarrow [0,1) \\ x = (\epsilon_1, \ldots) &\longmapsto 2\sum_{j=1}^{\infty} \epsilon_j 3^{-j} \end{aligned}$$

on the commutative group on W, taken without a topology. Then, if we provide W with the topology from the range of J_3, so that J_3 is a homeomorphism onto the $\frac{1}{3}$ Cantor set, then W is a compact, totally disconnected group.

d. In *Definition 5.3.11*, we defined the Walsh functions on $[0,1)$. *The Walsh functions can be identified with* $\widehat{W}$ by means of J in the following way.

First note that since W is compact, with the product topology, then $\widehat{W}$ is a discrete commutative group [Rud62, Section 1.7.3], i.e., a LCAG with the discrete topology. Now, for each $x \in W$, $x = (\epsilon_1, \ldots)$, $\epsilon_j \in \{0,1\}$, and for each $n \in \mathbb{N}$, we define

$$(44)\qquad \begin{aligned} \rho_n : W &\longrightarrow \{+1, 1\} \\ x &\longmapsto 1 - 2\epsilon_n. \end{aligned}$$

$$
\begin{tikzcd}
W \arrow[rr, "J"] \arrow[dr, "\rho_n"'] & & {[0,1)} \arrow[dl, "r_n"] \\
& \{+1,-1\} &
\end{tikzcd}
$$

Figure 27:

$$
\begin{tikzcd}
W \arrow[rr, "J"] \arrow[dr, "\omega_n"'] & & {[0,1)} \arrow[dl, "w_n"] \\
& \{+1,-1\} &
\end{tikzcd}
$$

Figure 28:

It is easy to see that the mapping J allows us to identify the sequence $\{\rho_n\}$ defined by (44) with the sequence of Rademacher functions defined in *Definition 5.3.11* a.

For example, if $n = 1$, then $Jx \in [0, \frac{1}{2})$ in the case $\epsilon_1 = 0$ and $Jx \in [\frac{1}{2}, 1)$ in the case $\epsilon_1 = 1$. Thus, since $\rho_1((0, \epsilon_2, \epsilon_3, \ldots)) = 1 - 0 = 1$ and $\rho_1((1, \epsilon_2, \epsilon_3, \ldots)) = 1 - 2 \cdot 1 = -1$, we have $\rho_1 = r_1 \circ J$, see *Figure 1.*

Note that each function ρ_n is an element of $\widehat{W}$, see Exercise 5.41

Next, let $\omega_0 : W \rightarrow \{+1, -1\}$ be identically 1, and for each $n \in N$, define

$$
\begin{aligned}
\omega_n : W &\longrightarrow \{+1, -1\} \\
x &\longmapsto \prod_{\epsilon_j = 1} (x, \rho_j),
\end{aligned}
\tag{45}
$$

where $n = \sum_{j=1}^{\infty} \epsilon_j 2^{j-1}$ is the binary expansion of n. It is easy to see that the mapping J allows us to identify the sequence $\{\omega_n : n \geq 0\}$ with the sequence of Walsh functions defined in *Definition 5.3.11* c, see *Figure 3*, see Exercise 5.41.

#53 Fall 1997

Further, each function ω_n is an element of $\widehat{W}$.

We parameterize $\{\omega_n\}$ in the following way. For any level r, let $W^r \subseteq W$ be the subgroup of W consisting of elements $x = (\epsilon_1, \ldots)$ with the property that $\epsilon_{r+1} = \epsilon_{r+2} = \cdots = 0$. Clearly, $W^r \subseteq W^{r+1}$ for all r. Also, if $n = \sum_{j=1}^{r} \epsilon'_j 2^{j-1}$, then $x = (\epsilon'_1, \ldots, \epsilon'_r, 0, 0, \ldots) \in W^r$. Further, for this n, ω_n is associated with X^r_n in the tree $\{X^r_m : r \geq 0 \text{ and } m = \sum_{j=1}^{r} \epsilon_j 2^{j-1}\}$ taken with bit reversal ordering; and, in fact, ω_n is associated with X^{r+q}_n for all $q \in \mathbb{N}$, where $n = \sum_{j=1}^{r+q} \epsilon'_j 2^{j-1}$ and $\epsilon'_j = 0$ for $r + 1 \leq j \leq r + q$. Formulate an exercise from this note.

Investigate the discrete group $\widehat{W}$ more fully.

e.

- **Do p-adic number and prove that the Walsh (dyadic) group is the ring of integers for the 2-adic numbers. See my old preprints on algebraic number fields.**
- **Develop theory of MRA-groups beginning with Vilenkin groups.**
- **Do atomic Hardy spaces $0 < p \leq 1$. We have**

$$H^p(G) = L^p(G) \textbf{ if } p > 1$$
$$H^1(G) \subseteq L^1(G)$$

  where G is a Vilenkin group. In particular, take $G = W$.
- **Look at subgroups of W, e.g., let $G_n = \{(0, 0, \ldots, 0, \epsilon_{n+1}, \epsilon_{n+2}, \ldots)\} \subseteq W$. G_n corresponds to $[0, \frac{1}{2^n})$. In fact, $\frac{1}{2^{n+1}} + \frac{1}{2^{n+2}} + \cdots = \frac{1}{2^{n+1}}(1 + \frac{1}{2} + \frac{1}{2^2} + \ldots) = \frac{1}{2^n}$. Compare with $\{W^r\}$.**
- **See work of Daly and Phillips, e.g., [DP98].**

Definition 5.3.14 (Wavelet Packets)

We defined the sequence of wavelet packet filters in Definition 5.2.11. For a given MRA these were the sequence of products

$$\prod_{j=1}^{r} H_{\epsilon_j}(2^j \gamma).$$

Should I put the actual definition of wavelet packets (that I'm about to give) in Section 5.2?

a. Let $H_0, H_1 \in L^2(\mathbb{T})$, and let $\{h_0[n]\}$, $\{h_1[n]\}$ be the sequence of Fourier coefficients of H_0, H_1, respectively.

A sequence $\{w_i : i = 0, 1, \ldots\} \subseteq L^2(\mathbb{R})$ of functions is a sequence of *wavelet packets* corresponding to H_0, H_1 if

$$w_{2i}(t) = \sum_n h_0[n]\sqrt{2}\, w_i(2t - n) \text{ in } L^2(\mathbb{R}) \tag{46}$$

and

$$w_{2i+1}(t) = \sum_n h_1[n]\sqrt{2}\, w_i(2t - n) \text{ in } L^2(\mathbb{R}) \tag{47}$$

for all $i = 0, 1, \ldots$. In this case, the sequence $\{w_{i,j,k} : i = 0, 1, \ldots,\ j, k \in \mathbb{Z}\}$, defined by

$$w_{i,j,k}(t) = 2^{\frac{j}{2}} w_i(2^j t - k),$$

is a sequence of *generalized wavelet packets* corresponding to H_0, H_1.

b. The definition in part a can be used in the following special cases or settings:

   i. H_0, H_1 are QMFs;

   ii. H_0, H_1 are QMFs and $h_1[n] = (-1)^n \overline{h_0[-n+1]}$;

   iii. H_0, H_1 are QMFs arising from an MRA $\{V_j\}$ with scaling function ϕ and $h_1[n] = (-1)^n \overline{h_0[-n+1]}$.

   Further, it is often reasonable to deal with $2N$-tap QMFs H_0, H_1, and $\operatorname{supp} w_i \subseteq [0, 2N-1]$.

Make sure I have defined $2N$-tap and/or FIR filter.

Show the equivalence of the procedure in Section 5.2.10 and in this definition. Definitely do this.

Should I talk about sequency now that we've defined $w_{i,j,k}$?

Example 5.3.15 (The Walsh Functions as Waveletpackets)

a. Consider the Haar MRA from Example 5.2.7. Then the scaling function is $\phi = \mathbf{1}_{[0,1)}$, and the Fourier coefficients of the QMF H_0 are $h_0[0] = h_0[1] = \frac{1}{\sqrt{2}}$ and $h_0[n] = 0$ otherwise. Thus,

$$H_0(\gamma) = \frac{1}{\sqrt{2}}(1 + e^{-2\pi i\gamma}) = \sqrt{2}\, e^{-\pi i\gamma} \cos \pi\gamma. \tag{48}$$

   Using the MRA algorithm (in particular, $h_1[n] \equiv (-1)^n \overline{h_0[-n+1]}$), we showed that $h_1[0] = \frac{1}{\sqrt{2}}$, $h_1[1] = \frac{-1}{\sqrt{2}}$, and $h_1[n] = 0$ otherwise. Thus,

$$H_1(\gamma) = \frac{1}{\sqrt{2}}(1 - e^{-2\pi i\gamma}) = \sqrt{2}\, ie^{-\pi i\gamma} \sin \pi\gamma;$$

   and we computed that the orthonormal wavelet is

$$\psi = \mathbf{1}_{[0,\frac{1}{2})} - \mathbf{1}_{[\frac{1}{2},1)}.$$

b. Using Definition 5.3.13 (Equations 46 and 47), we shall compute the wavelet packets corresponding to the Haar MRA. We begin with $j = 0$. Then

$$w_0(t) = w_0(2t) + w_0(2t-1) \tag{49}$$

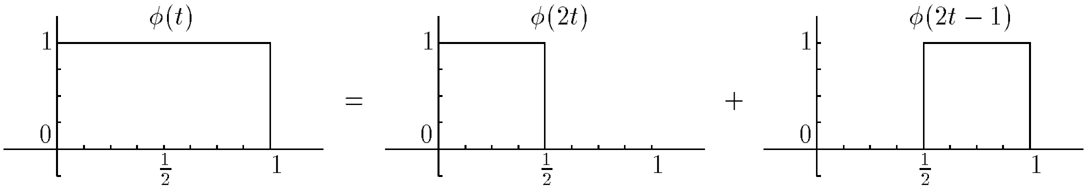


Figure 29:

and

$$w_1(t) = w_0(2t) - w_0(2t-1). \tag{50}$$

Note that $w_0 = \phi$ and $w_1 = \psi$ are solutions of 49 and 50, respectively, see *Figure 2*.

Discuss uniqueness questions concerning scaling functions. Insert my little result on the characterization of scaling functions. See thw work of Colella, Heil et al. on scaling equations, including [Dau92].

Since $2t-1 \in [0,1) \Rightarrow 2t \in [1,2) \Rightarrow t \in [\frac{1}{2}, 1)$, **we see that** $\phi(2t-1)$ **is depicted as in** ***Figure 2*.**

c. The real problem, of course, is to *solve* for w_0 in Equation 49. Formally, the Fourier transform of Equation 49 is

$$\widehat{w}_0(\gamma) = \frac{1}{2}(1 + e^{-\pi i \gamma})\widehat{w}_0(\frac{\gamma}{2}). \tag{51}$$

Repeating Equation 51 n-times and noting the definition of H_0 in this case of the Haar MRA (Equation 5.X), we have

$$\widehat{w}_0(\frac{\gamma}{2}) = \frac{1}{2}(1 + e^{-\pi i \frac{\gamma}{2}})\widehat{w}_0(\frac{\gamma}{4}) = \frac{1}{\sqrt{2}} H_0(\frac{\gamma}{4})\widehat{w}_0(\frac{\gamma}{4})$$

$$
\begin{aligned}
\widehat{w}_0(\gamma) &= \frac{1}{\sqrt{2}} H_0(\frac{\gamma}{2}) \widehat{w}_0(\frac{\gamma}{2}) \\
&= \frac{1}{\sqrt{2}} H_0(\frac{\gamma}{2}) \frac{1}{\sqrt{2}} H_0(\frac{\gamma}{2^2}) \widehat{w}_0(\frac{\gamma}{2^2}) \\
&= \prod_{j=1}^{n} (\frac{1}{\sqrt{2}} H_0(\frac{\gamma}{2^j})) \widehat{w}_0(\frac{\gamma}{2^n}) \\
&= \widehat{w}_0(\frac{\gamma}{2^n}) \prod_{j=1}^{n} e^{-\pi i \frac{\gamma}{2^j}} \cos \pi \frac{\gamma}{2^j}.
\end{aligned} \tag{52}
$$

Suppose we ask precisely for solution $w_0 \in L^1(\mathbb{R})$ of (49) which are normalized so that $\int w_0(t)dt = 1$. Then because of Viète's formula (Exercise 1.36) and the facts that $\widehat{w}$ is continuous and $\widehat{w}(0) = 1$, we have from (52) that

$$
\begin{aligned}
\widehat{w}_0(\gamma) &= exp\{-\pi i \sum_{j=1}^{\infty} \frac{1}{2^j}\} \prod_{j=1}^{\infty} \cos \pi \frac{\gamma}{2^j} \\
&= e^{-\pi i \gamma} \frac{\sin \pi\gamma}{\pi\gamma}.
\end{aligned}
$$

Thus, we have *proved* that $w_0 = \phi$ is the only solution of (49) from the class $L^1(\mathbb{R})$ with the constraint $\widehat{w}_0(0) = 1$.

| **Can you think of any other, perhaps distributional, solutions of (49)?**

Note that $w_1 = \psi$ is the unique solution of (50) given that $w_0 = \phi$.

d. Now that we have $w_0 = \phi$ and $w_1 = \psi$ we go back to Equations 46 and 47 to solve for the other wavelet packets corresponding to the Haar MRA.

If $j = 1$, then

$$
w_2(t) = w_1(2t) + w_1(2t-1) \tag{53}
$$

and

$$
w_3(t) = w_1(2t) - w_1(2t-1). \tag{54}
$$

The functions $w_1(2t)$ and $w_1(2t-1)$ are illustrated in *Figure 4*.

The function w_2 and w_3 are easily computed and are depicted in *Figure 4*.

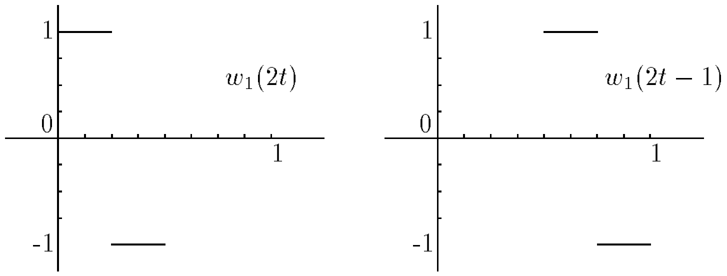


Figure 30: Translated Dilates of w_1.

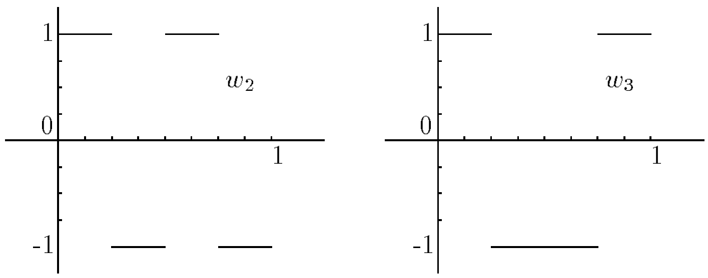


Figure 31: w_2 and w_3

e. The wavelet packets $\{w_n\}$ are defined on $\mathbb{R}$, and the ones we have exhibited are the Walsh functions $\{w_n\}$, when the latter which we have defined on $\mathbb{T}$ are adjusted to vanish outside of $[0, 1)$.

Prove that the wavelet packets $\{w_n\}$ corresponding to the Haar MRA are precisely the Walsh functions, where the latter are adjusted (as above) to be defined as functions on $\mathbb{R}$ supported by $[0, 1]$.

- **Reference Meyer's book (translated by Ryan) as well as the CMW article in Ruskai et al.**
- **Quote general wavelet packet theorems.**
- **Tie in the sequencing parameter i in $\{w_{i,j,k}\}$ with frequency localization.**
- **Give special student project on sequency.**

5.4 Bessel Sequences

Definition 5.4.1 (Bessel and Hilbert Sequences)
Let H, H_1, and H_2 be Hilbert spaces.

a. $L \in \mathcal{L}(H_1, H_2)$ is a *topological isomorphism* if it is bijective, in which case $L^{-1} \in \mathfrak{L}(H_2, H_1)$ by the Open Mapping Theorem (Theorem B.9).

b. A sequence $\{x_n : n \in \mathbf{Z}^d\} \subseteq H$ is a *Bessel sequence* if the *Bessel mapping*

$$L : H \mapsto l^2(\mathbf{Z}^d),$$
$$y \mapsto \{\langle y, x_n\rangle\}$$

is a well-defined linear function.

c. A sequence $\{x_n : n \in \mathbf{Z}^d\} \subseteq H$ is a *Hilbert sequence* if

$$\forall c = \{c[n]\} \in l^2(\mathbf{Z}^d), \exists y_c \in H \text{ such that}$$
$$\forall n \in \mathbf{Z}^d, c[n] = \langle y_c, x_n\rangle.$$

d. Clearly, a Bessel sequence is a Hilbert sequence if and only if the Bessel mapping L is surjective.

Theorem 5.4.2 Continuity of L and Reprensentation of L* for Bessel Mappings
Let H be a Hilbert space, and let $\{x_n : n \in \mathbf{Z}^d\}$ be a Bessel sequence with Bessel mapping L.

a. L is continuous, i.e.,

$$\exists B > 0 \quad \text{such that} \quad \forall y \in H, \ \sum |\langle y, x_n\rangle|^2 \leq B\, \|y\|^2 . \tag{55}$$

b. For each $c = \{c[n]\} \in l^2(\mathbf{Z}^d)$,

$$L^* c = \sum c[n] x_n \text{ in } H. \tag{56}$$

Proof

a. We first define the linear maps

$$L_N : H \to l^2(\mathbf{Z}^d)$$
$$y \mapsto \{\langle y, x_n\rangle : |n| \le N\},$$

where $n = (n_1, ..., n_d)$ and $|n|^2 = \sum_{j=1}^d |n_j|^2$, (see B). Each L_N is obviously continuous.

Since $L_y \in l^2(\mathbf{Z}^d)$ for $y \in H$, we have

$$||L_N y - Ly||^2_{l^2(\mathbf{Z}^d)} \quad = \quad \sum_{|n|>N} |\langle y, x_n\rangle|^2 ,$$

and the right side goes to 0 as $N \to \infty$. Thus,

$$\forall y \in H, \quad \lim_N L_N y \quad = \quad Ly \quad \text{in } l^2(\mathbf{Z}^d);$$

and so L is continuous by the Uniform Boundedness Principle (Theorem B.8). In fact,

$$\sup_N ||L_N|| \quad = \quad B^{\frac{1}{2}} \quad < \quad \infty.$$

b. To prove (5.4.2) we use (5.4.1) in the following way. Let $c \in l^2(\mathbf{Z}^d)$ and let

$$s_N \quad = \quad \sum_{|n|\le N} c[n]x_n.$$

Then for $N > M$, the definition of the norm and Hölder's Inequality allow us to write

$$||s_N - s_M||^2 \quad = \quad \sup_{||y||=1} | \sum_{M<|n|\le N} c[n]\langle x_n, y\rangle \ |^2$$
$$\le \sup_{||y||=1} \sum_{M<|n|\le N} |c[n]|^2 \sum_{M<|n|\le N} |\langle x_n, y\rangle|^2$$
$$\le B \sup_{||y||=1} ||y||^2 \sum_{M<|n|\le N} |c[n]|^2.$$

Therefore $\{s_N\} \subseteq H$ is a Cauchy sequence and so $\sum c[n]x_n$ converges in H. Let $c = \{c[n]\} \in l^2(\mathbf{Z}^d)$ and take $x \in H$. Then

$$\langle x, L^*c\rangle = \langle Lx, c\rangle = \sum (Lx)[n]\overline{c[n]} \tag{57}$$

$$= \sum \langle x, x_n \rangle \overline{c[n]} = \sum \langle x, c[n]x_n \rangle.$$

Since $\sum c[n]x_n$ converges in H we can use the continuity of the inner product (Proposition B.) to write the right side of (5.4.2) as $\langle x, \sum c[n]x_n \rangle$. Thus,

$$\forall x \in H, \ \ \langle x, L^* c \rangle = \langle x, \sum c[n]x_n \rangle, \tag{58}$$

and so (5.4.4) yields (5.4.2).

Figure 5.57

The mapping L is a *digitization* of H. It provides a discrete set with which to carry out a signal processing of signals y from H. Even if H is discrete and/or finite, $L(H)$ may be a more manageable discrete set than H itself. Although not necessarily consisting of sampled values of y, Ly is modeled as accessible data from which we reconstruct y.

| **Put Noise Reduction Remarks For This Figure Later On**

Definition 5.4.3 (Biorthonormal and Minimal Sequences)

Let H be a Hilbert space. A sequence $\{x_n : n \in \mathbf{Z}^d\} \subseteq H$ is *minimal* if

$$\forall n \in l^2(\mathbf{Z}^d), x_n \notin span\{x_k : k \neq n\}.$$

Two sequences $\{x_n : n \in \mathbf{Z}^d\}$ and $\{y_n : n \in \mathbf{Z}^d\}$ are *biorthonormal* if

$$\forall m, n \in \mathbf{Z}^d, \langle x_m, y_n \rangle = \delta(m, n).$$

| **Make note about biorthogonal wavelets.**

Proposition 5.4.4

Let H be a Hilbert space and let $\{x_n : n \in \mathbf{Z}^d\} \subseteq H$.

a. There is a sequence $\{y_n : n \in \mathbf{Z}^d\}$ such that $\{x_n\}$ and $\{y_n\}$ are biorthonormal if and only if $\{x_n\}$ is minimal.

b. Furthermore, $\{y_n\}$ of part a is uniquely determined if and only if $\{x_n\}$ is not only minimal but also $\overline{span}\{x_n\} = H$.

Proof

a. Let $\{x_n\}$ and $\{y_n\}$ be be biorthonormal, and let $X_n = \overline{span}\{x_k : k \neq n\}$. Then y_n is orthogonal to X_n. On the other hand $\langle x_n, y_n \rangle = 1$, and so $x_n \notin X_n$.
If $\{x_n\}$ is minimal then $x_n \notin X_n$. Thus, by the Hahn-Banach Theorem there is $y_n \in H$ such that $\langle x_n, y_n \rangle = 1$ and y_n is orthogonal to X_n. Hence, $\{x_n\}$ and $\{y_n\}$ are biorthonormal.

b. Part *b* is clear by the Hahn-Banach Theorem.

Do Details

Remark 5.4.5 (Stationarity and Biorthonormality)

a. If H is a Hilbert space, $\{x_n : n \in \mathbf{Z}^d\} \subseteq H$ is minimal and $\overline{span}\{x_n\} = H$, and $\{x_n\}$ and $\{y_n\}$ are biorthonormal, then it is not necessarily true that $\overline{span}\{y_n\} = H$. An example of Kaczmarz and Steinhaus (1935) provides a counterexample. In fact, let $\{u_n : n \in \mathbb{N}\}$ be an ONB for H, and define

$$\forall n \in \mathbb{N}, \quad x_n = u_1 + u_{n+1} \quad \text{and} \quad y_n = u_{n+1}.$$

Clearly, $\{x_n\}$ and $\{y_n\}$ are biorthonormal. However, $\overline{span}\{y_n\} \neq H$ since u_1 is orthogonal to $\{y_n\}$ and since $\{u_n\}$ is an ONB. To see that $\overline{span}\{x_n\} = H$, suppose $\langle y, x_n \rangle = 0$ for all n, i.e.,

$$\forall n \in \mathbb{N}, \quad \langle y, u_1 \rangle = -\langle y, u_{n+1} \rangle \tag{59}$$

By Theorem B.a , $\{\langle y, x_{n+1} \rangle\} \in l^2(\mathbb{N})$, and so $\lim_{n\to\infty} \langle y, x_{n+1} \rangle = 0$. Thus, $\langle y, u_1 \rangle = 0$, which by (5.4.5), also implies that $\langle y, u_n \rangle = 0$ for all n. Since $\overline{span}\{u_n\} = H$, we have $y = 0$; and, therefore, by our initial supposition, $\overline{span}\{x_n\} = H$.

b. On the other hand, Masani and Rozanov[Mas 60] proved that $\overline{span}\{y_n : n \in \mathbf{Z}^d\} = H$ in the case that $\{x_n : n \in \mathbf{Z}^d\}$ is not only minimal and $\overline{span}\{x_n\} = H$, but also *stationary*, i.e., $\{\langle x_{n+k}, x_k \rangle : k, n \in \mathbf{Z}^d\}$ is independent of k. Stationary sequences were defined in Definition 3.7.3. The proof that $\overline{span}\{y_n\} = H$ uses Kolmogorov's mapping discussed in Remark 3.7.4.

Where should I insert stationarity of Gabor and nonstationarity of wavelets? See W:M & A Ch 7 for details. Probably insert here. Tie-in with biorthogonal wavelets $\langle \psi_{m,n}, \tilde{\psi}_{m,n} \rangle$ where we have non-stationarity & $\overline{span}\{\psi_{m,n}\} = H$.

In Section 3.7 do I do $\lim \frac{1}{2N+1} \sum_{|k|\leq N} x_{n+k}\overline{x_k} = P[n]$? If not, mention it now.

Definition 5.4.6 (Uniformly Discrete Sequences)

A sequence $\{t_n : n \in \mathbf{Z}\} \subseteq \mathbb{R}$ is *uniformly discrete* with separation r if $\inf\{|t_m - t_n| : m \neq n\} = r > 0$.

Theorem 5.4.7 Plancherel-Pólya Theorem

Let $f \in PW_\Omega$ and let $\{t_n\}$ be uniformly discrete with separation r. Then

$$\sum |f(t_n)|^2 \leq B\, \|f\|^2_{L^2(\mathbb{R})} \tag{60}$$

where $B > 0$ can be taken to be

$$B = \Big(\frac{2}{\pi r}\Big)^2 \frac{1}{2\Omega}(e^{2\pi\Omega r} - 1)$$

Proof

a. The first step, which is classical and which we shall omit, involves a Phragmén-Lindelöf argument and asserts that

$$\int |f(t+iy)|^2 dt \leq e^{4\pi\Omega|y|}\, \|f\|^2_{L^2(\mathbb{R})}, \tag{61}$$

e.g., [PP37], [Pla 49], [You 80, pages 94-96], cf., Remark 5.4.8 for a different formal proof which can be verified and generalized.

b. Let $f \in PW_\Omega$. Then

$$\forall z_0 \in \mathbb{C} \text{ and } \forall \rho > 0,\ \ |f(z_0)|^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi} \left|f(z_0 + \varrho e^{i\theta}\right|^2 d\theta, \tag{62}$$

since f^2 is entire, cf., [Ben 92b, pages 474-475] for a discussion of (5.4.8) in terms of subharmonicity because of our proof of the analogue of Theorem 5.4.7 in $\mathbb{R}^d$.

c. We multiply both sides of (5.4.8) by ϱ and integrate between 0 and $\frac{r}{2}$ to obtain

$$\forall z_0 \in \mathbb{C}, \quad |f(z_0)|^2 \leq \frac{4}{\pi r^2}\int_0^{\frac{r}{2}}\int_{-\pi}^{\pi} \varrho \left|f(z_0 + \varrho e^{i\theta})\right|^2 d\theta\, d\varrho$$

$$= \frac{4}{\pi r^2}\iint_{|z-z_0|\leq \frac{r}{2}} |f(z)|^2 dt\,dy, \ \ z = t + iy.$$

Thus,

$$\sum |f(t_n)|^2 \leq \frac{4}{\pi r^2}\sum \iint_{|z|\leq\frac{r}{2}} |f(z+t_n)|^2 dt\,dy$$

$$\leq \frac{4}{\pi r^2}\int_{\frac{-r}{2}}^{\frac{r}{2}} \sum \int_{\frac{-r}{2}}^{\frac{r}{2}} |f(t + t_n + iy)|^2 dt\,dy$$

$$= \frac{4}{\pi r^2} \int_{\frac{-r}{2}}^{\frac{r}{2}} \sum \int_{t_n - \frac{r}{2}}^{t_n + \frac{r}{2}} |f(t+iy)|^2 dt dy$$

$$\leq \frac{4}{\pi r^2} \int_{\frac{-r}{2}}^{\frac{r}{2}} \int |f(t+iy)|^2 dt dy, \tag{63}$$

where the last inequality follows by the definition of r.

d. Combining (5.4.7) with (5.4.9) we obtain

$$\sum |f(t_n)|^2 \leq \frac{4}{\pi r^2} \int_{\frac{-r}{2}}^{\frac{r}{2}} e^{4\pi\Omega|y|} \, \|f\|^2_{L^2(\mathbb{R})} \, dy$$

which gives the result. □

$$\frac{4}{\pi r^2} \int_{\frac{-r}{2}}^{\frac{r}{2}} e^{4\pi\Omega|y|} \, \|f\|^2_{L^2(\mathbb{R})} \, dy$$

$$= \frac{8}{\pi r^2} \int_{0}^{\frac{r}{2}} e^{4\pi\Omega|y|} \, \|f\|^2_{L^2(\mathbb{R})} \, dy$$

$$= (\frac{2}{\pi r^2})^2 \frac{2}{4\Omega} (e^{4\pi\Omega\frac{r}{2}} - 1) \, \|f\|^2_{L^2(\mathbb{R})} \, .$$

Remark 5.4.8 (Bessel Sequences and the Plancherel-Pólya Theorem)

a. Let $H = PW_\Omega$. Theorem 5.4.7 asserts that *if* t_n *is uniformly discrete then*

$$\{x_n(t)\} = \{\tau_{t_n} d_{2\pi\Omega}\} \subseteq PW_\Omega \tag{64}$$

is a Bessel sequence with Bessel mapping

$$L : PW_\Omega \to l^2(\mathbb{Z})$$

$$f \mapsto \{f(t_n)\}.$$

To verify this claim we need only note that

$$\widehat{x}_n(\gamma) = e^{-2\pi i t_n \gamma} \mathbf{1}_{[-\Omega,\Omega]}(\gamma)$$

and

$$f(t_n) = \int_{-\Omega}^{\Omega} \widehat{f}(\gamma) e^{2\pi i t_n \gamma} d\gamma = \langle f, x_n \rangle.$$

| **See "Proof" Of This Form Of Inversion In Exercise 1.51**

b. The following formal calculation for $f \in PW_\Omega$ can be made rigorous, not only to provide an alternative to the Phragmén-Lindelöf step in Theorem 5.4.7 , but to give a means for dealing with analogous problems in $\mathbb{R}^d$:

$$\int |f(t+iy)|^2 dt = \int \left| \int_{-\Omega}^{\Omega} \hat{f}(\gamma) e^{2\pi i t\gamma} e^{-2\pi y\gamma} d\gamma \right|^2 dt$$

$$= \int_{-\Omega}^{\Omega} \int_{-\Omega}^{\Omega} \hat{f}(\gamma)\overline{\widehat{f}(\lambda)} e^{-2\pi y(\gamma+\lambda)} \left(\int e^{2\pi i t(\gamma-\lambda)} dt \right) d\lambda d\gamma$$

$$= \int_{-\Omega}^{\Omega} \overline{\widehat{f}(\lambda)} e^{-2\pi y\lambda} \int_{-\Omega}^{\Omega} \delta(\gamma-\lambda)\widehat{f}(\gamma) e^{-2\pi y\gamma} d\gamma d\lambda$$

$$= \int_{-\Omega}^{\Omega} \left|\hat{f}(\lambda)\right|^2 e^{-4\pi y\lambda} d\lambda \leq e^{4\pi|y|\Omega} \, ||f||^2_{L^2(\mathbb{R})},$$

where the manipulations with δ and $\int e^{2\pi i t(\gamma-\lambda)} dt$ can be dealt with properly by approximate identity arguments and the methods of Chapter 2.

c. Boas[Boa 40, Theorems 1,2,3] gave a proof of Theorem 5.4.7 different from the argument of Plancherel and Pólya. An attractive feature of Boas' approach is that it leads us to write (5.4.6) as a *weighted Fourier transform norm inequality*, viz.,

$$||f||^2_{L^2{}_\nu(\mathbb{R})} \leq B||f||^2_{L^2{}_\mu(\hat{\mathbb{R}})}$$

In the case of (5.4.6), ν is the discrete measure $\sum \delta_{t_n}$, μ is the absolutely continuous measure $\mu = \mathbf{1}_{[-\Omega,\Omega]}$ (usually more comfortably written as "$d\mu(\gamma) = \mathbf{1}_{[-\Omega,\Omega]}(\gamma)d\gamma$"), and $L^2_\mu(\hat{\mathbb{R}})$ is $(PW_\Omega)^\wedge$. Weights such as $\mu = \mathbf{1}_{[-\Omega,\Omega]}$ play a role in the Bell Labs uncertainty principle theory from the early 1960s ([Pric 85, pages 201-220 by H.J.Landau]) ; and other weighted L^p-spaces have applications in prediction theory (Remark 3.7.4), restriction theory [Ste 93], and general uncertainty principle inequalities. Weighted inequalities generalizing (5.4.6) are found in [BF94,Chapter 7].

d. Boas and Plancherel-Pólya are also responsible for proving "opposite inequalities" to (5.4.6), e.g., [Boa 54, pages 196-199]. This, along with work of Paley and Wiener[PW34] and Levinson[Lev 40] are a prelude to the theory of frames (Section ??).