UNIVERSIDAD NACIONAL DE CÓRDOBA

**FACULTAD DE MATEMÁTICA, ASTRONOMÍA Y FÍSICA**

SERIE "B"

# TRABAJOS DE MATEMATICA

**Nº 57/2010**

## Short Course on Wavelets

Edward N. Wilson

Editores: Jorge R. Lauret–Jorge G. Adrover

CIUDAD UNIVERSITARIA – 5000 CÓRDOBA

REPÚBLICA ARGENTINA

# University of Cordoba, Argentina

## Short Course on Wavelets: Aug. 19-Aug. 23, 2010

### Speaker: Edward N. Wilson
### Washington University in St. Louis

## Topics covered in the Short Course

1. Definitions of basic unitary operators on $L^2(\mathbb{R}^n)$ and their properties.

2. Gabor, Haar, and Shannon Reproducing Function Systems

3. Shift Invariant Spaces

4. Applications of Shift Invariant Space Theory to Wavelet Systems

5. Continuous Wavelet Transforms and the Unified Theorem

## 1. Definitions of basic unitary operators on $L^2(\mathbb{R}^n)$ and their properties

**1.1 Terminology.** In signal processing, the real line $\mathbb{R}$ is callled the time domain while a copy $\widehat{\mathbb{R}}$ of $\mathbb{R}$ is called the frequency domain. For practical applications, a one-dimensional signal is a square integrable function $f$ from $\mathbb{R}$ to $\mathbb{R}$ which is at least piecewise continuous. For the mathematical theory of signal processing, it's more convenient to regard signals as members of the Hilbert space $L^2(\mathbb{R})=\{a.e.$ equivalence classes of measurable functions from $\mathbb{R}$ into $\mathbb{C}$ which are square integrable with respect to Lebesgue measure$\}$.

When we move from $\mathbb{R}$ to $\mathbb{R}^n$, $n \geq 2$, and discuss higher dimensional mathematical signals $f \in L^2(\mathbb{R}^n)$, it's convenient as in calculus classes to take $\mathbb{R}^n$ to be the space of $1 \times n$ real column matrices $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ and call it the space-domain. We take $(\mathbb{R}^n)^\frown$ to be the space of $n \times 1$ real row matrices $\xi = [\xi_1 \quad \xi_2 \cdots \quad \xi_{n-1} \quad \xi_n]$ and call it the $n-$ dimensional frequency domain. These two spaces play very different roles so "identifying them" by replacing both column and row matrices with ordered $n-$ tuples is usually not a good idea. Note that the matrix product $\xi x$ of $\xi \in (\mathbb{R}^n)^\frown$ and $x \in \mathbb{R}^n$ is just $\sum_{j=1}^{n} \xi_j x_j$, which in many calculus books is called the "dot product" of $n$-tuples. We write $dx$ and $d\xi$ for the increments of Lebesgue measure on $\mathbb{R}^n$ and $(\mathbb{R}^n)^\frown$.

**1.2 Definitions.** The basic unitary operators for harmonic analysis on $L^2(\mathbb{R}^n) = L^2(\mathbb{R}^n, dx)$ are as follows:

(*i*) Translation operators $(T_y f)(x) \equiv f(x-y)$, $y \in \mathbb{R}^n$;

(*ii*) Modulation operators $(M_\xi f)(x) \equiv e_\xi(x) f(x)$ where, for $\xi \in (\mathbb{R}^n)\hat{}$, $e_\xi(x) = e_x(\xi) \equiv e^{2\pi i \xi x}$;

(*iii*) Dilation operators $(D_a f)(x) \equiv |det\, a|^{-1/2} f(a^{-1}x)$, $a \in GL(n, \mathbb{R})$ = group of invertible $n \times n$ real matrices;

(*iv*) The Fourier transform $\mathcal{F} \equiv$ unique extension to a unitary operator from $L^2(\mathbb{R}^n)$ onto $L^2((\mathbb{R}^n)\hat{})$of the operator $f \mapsto \widehat{f}$ from $L^1(\mathbb{R}^n) \cap L^2((\mathbb{R}^n)\hat{})$ into {bounded, continuous functions on $(\mathbb{R}^n)\hat{}$ } defined by

$$\widehat{f}(\xi) = \int_{\mathbb{R}} f(x) e_{-\xi}(x) dx \tag{1}$$

*Warning: Some authors delete the $2\pi$ factor in the definition of the elementary exponentials $e_\xi$ and are then obliged to multiple the integral in (1) by $\frac{1}{(2\pi)^{n/2}}$ in order to make $\mathcal{F}$ unitary. Also, some authors denote the operator in (iii) by $D_{a^{-1}}$.*

**1.3 Notations**: (*i*) When $\mathcal{U}$ is an operator on $L^2(\mathbb{R}^n)$, $\widehat{\mathcal{U}}$ is the operator $\mathcal{F} \circ \mathcal{U} \circ \mathcal{F}^{-1}$ on $L^2((\mathbb{R}^n)\hat{}\ )$. Thus, for each $f \in L^2(\mathbb{R}^n)$, $\widehat{\mathcal{U}}\widehat{f} = (\mathcal{U}f)^\wedge$.

(*ii*) For $f \in L^2(\mathbb{R}^n)$, the support of $f$, denoted $supp\, f$, is the $a.e.$ well defined set $\{x \in \mathbb{R}^n : f(x) \neq 0\}$.

(*iii*) For S a measurable subset of $\mathbb{R}^n$, $L^2(S)$ denotes the subspace of $L^2(\mathbb{R}^n)$ consisting of all $f \in L^2(\mathbb{R}^n)$ for which, modulo a Lebesgue null set, $supp\, f \subset S$ or, equivalently, $f = \chi_S f\, a.e.$ Here $\chi_S$ denotes the characteristic function of S, *i.e.* the function which is 1 on S and is 0 off S.

**1.4 Theorem on the Geometric/Algebraic Properties of the Basic Operators.**

$(i)$ On the Hilbert space $L^2(\mathbb{R}^n)$ :

$y \mapsto T_y$ is a unitary representation of the additive group $(\mathbb{R}^n, +)$ (thus, $T_x \circ T_y = T_{x+y}$ for all $x, y$ and $x \mapsto T_x f$ is continuous from $\mathbb{R}^n$ into $L^2(\mathbb{R}^n)\, \forall f \in L^2(\mathbb{R}^n)$);

$\xi \mapsto M_\xi$ is a unitary representation of the additive group $(\mathbb{R}^n)\hat{}$ ;

$a \mapsto D_a$ is a unitary representation of the multiplicative group $GL(n, \mathbb{R})$ $(the\ determinant\ factor\ is\ what\ makes\ makes\ each$ $D_a\ unitary\ and\ multiplying\ x\ by\ a^{-1}\ in\ the\ definition\ of\ D_a is$ $what\ is\ needed\ to\ have\ D_{ab} = D_a D_b)$.

$(ii)$ For all $(a, y, \xi) \in (\mathbb{R}\backslash\{0\}) \times \mathbb{R}^n \times (\mathbb{R}^n)\hat{}$ :

$D_a T_y = T_{ay} D_{a;}$
$D_a M_\xi = M_{\xi a^{-1}} D_{a;}$
$M_\xi T_y = e_{+\xi}(y) T_y M_\xi;$
$(T_y)\hat{} = M_{-y};$
$(M_\xi)\hat{} = T_{\xi;}$
$((D_a)\hat{}\,\widehat{f}\,)(\xi) = |det\, a|^{1/2}\,\widehat{f}\,(\xi a)$ $(Multiply\ on\ the\ right\ by$ $a\ for\ (D_a)\hat{}\ instead\ of\ multiplying\ on\ the\ left\ by\ a^{-1}\ for\ D_a)$.

$(iii)$ When $S \subset \mathbb{R}^n$ is a bounded set:
$T_y$ maps $L^2(S)$ onto $L^2(S+y)$;
$M_\xi$ maps $L^2(S)$ onto itself;
$D_a$ maps $L^2(S)$ onto $L^2(aS)$ ($so\ supports\ get\ multiplied\ by$ $a,\ not\ a^{-1}$);

$\mathcal{F}$ maps $L^2(S) \subset L^1(\mathbb{R}^n)$ into the space of square integrable, real analytic functions on $(\mathbb{R}^n)\hat{}$ which vanish at $\infty$ (real analytic means described by a converging power series); in the case $n = 1$, for $f \in L^2(S) \setminus \{0\}$, $supp\,\hat{f} = \hat{\mathbb{R}} \setminus \{\text{countable set}\}$.

*Comments on the proof of Theorem 1.4.* The last statement in Theorem 1.4$(iii)$ is part of the classical Riemann-Lebesgue Lemma saying that Fourier transforms of $L^1$ functions are continuous functions vanishing at $\infty$. The statement about real analyticity follows from expressing $e_{-\xi}(x)$ as a power series and integrating (1) term-by-term. In fact, for $n = 1$, this power series converges everywhere on $\mathbb{C}$ and extends $\hat{f}$ to a complex analytic (or holomorphic) function on $\mathbb{C}$ which can't have more than countably many zeros. All of the other statements in $(i) - (iii)$ are just elementary computations using only the definitions. It's a very good exercise to do these computations since the relationships in Theorem 1.4 are used repeatedly in constructions of different types of reproducing function systems.

## 2. Gabor, Haar, and Shannon Reproducing Function Systems

**2.1 Definitions:** $(i)$ A signal $f \in L^2(\mathbb{R})$ is said to be band-limited if $f = \mathcal{F}^{-1}(F)$ for some F in $L^2(\hat{\mathbb{R}})$ with $supp$ F compact. Then the frequency band for $f$ is the length of the smallest interval $I$ for which $supp\,F \subset I$. As in the above comments, every band-limited function is real analytic, vanishes at $\infty$, and square integrable--we can regard the everywhere well defined, smooth function $f$ as a member of $L^2(\mathbb{R})$ by identifying it with its $a.e.$ equivalence class. Of course, we return from $f$ to F by applying $\mathcal{F}$

so $\widehat{f} = \mathrm{F}$ *a.e.* Note that the reason $f$ is so nice is that F is integrable and compactly supported: but, nice as it is, $f$ need not be integrable so we need the measure-theoretic extension of $\mathcal{F}$ to get back to the usually not nice F.

(*ii*) The standard unit interval in $\widehat{\mathbb{R}}$ is $\mathrm{I}_1 = [-1/2, 1/2]$ and the standard band limited space of functions on $\mathbb{R}$ is $\mathbb{B}_1(\mathbb{R})$=$\{ f$=$\mathcal{F}^{-1}\mathrm{F}$: $\mathrm{F} \in \mathrm{L}^2(\mathrm{I}_1)\}$. The Shannon sampling and scaling function $\phi_S$ is the member of $\mathbb{B}_1(\mathbb{R})$ whose Fourier transform is $\chi_{_{\mathrm{I}_1}}$. Thus,

$\phi_S(x) = \int_{-1/2}^{1/2} e_x(\xi)d\xi$ where $e_x(\xi) = e_\xi(x) = e^{2\pi i \xi x}$.

By elementary calculus, $\phi_S(0) = 1$ and, for $x \neq 0$, $\phi_S(x) = \frac{\sin(\pi x)}{\pi x}$. Another name for $\phi_S(x)$ is $\mathrm{sinc}(x)$.

**2.2 Theorem on the Properties of $\mathbb{B}_1(\mathbb{R})$ and the sinc function**

(*i*) $\{\mathrm{T}_k(\mathrm{sinc}): k \in \mathbb{Z}\}$ is an orthonormal basis of $\mathbb{B}_1(\mathbb{R})$;

(*ii*) $\mathrm{T}_y(\mathbb{B}_1(\mathbb{R})) = \mathbb{B}_1(\mathbb{R})\ \forall y \in \mathbb{R}$.

(*iii*) For each $f \in \mathbb{B}_1(\mathbb{R})$, $\sum_{k\in\mathbb{Z}} f(k)T_k(\mathrm{sinc})$ converges pointwise unconditionally to $f$ and also converges to $f$ in $\mathrm{L}^2(\mathbb{R})$. More generally, for each choice of $x_0 \in \mathbb{R}$, $\sum_{k\in\mathbb{Z}} f(x_0 + k)\,\mathrm{sinc}(x - (x_0 + k))$ converges unconditionally to $f(x)$ for each $x$. In particular, $f$ is uniquely determined by its values on $\mathbb{Z}$+$x_0$.

(*iv*) Each real-valued $f \in \mathbb{B}_1(\mathbb{R})$ has a unique extension to a holomorphic function on $\mathbb{C}$ satisfying $\sum_{k\in\mathbb{Z}} (f(z-k))^2 = \|f\|_2^2$ for every $z \in \mathbb{C}$.

**Proof.** $(i)$ follows from the Fourier series fact that $\{e_{-k}\chi_{_{I_1}} : k \in \mathbb{Z}\}$ is an orthonormal basis of $L^2(I_1)$ = image of $\mathbb{B}_1(\mathbb{R})$ under the unitary map $\mathcal{F}$ and use of Theorem 1.4 to see that $\mathcal{F}(T_k(\mathrm{sinc}))=e_{-k}\chi_{_{I_1}}$ for each $k \in \mathbb{Z}$.

$(ii)$ follows from Theorem 1.4 and the observation that $L^2(S)$ is invariant under the modulation operators $M_y$, $y \in \mathbb{R}$, for every measurable set $S \subset \widehat{\mathbb{R}}$.

$(iii)$ For each $f \in \mathbb{B}_1(\mathbb{R})$, we have $F=\widehat{f}$ in $L^2(I_1)$ and, for each $k \in \mathbb{Z}$, $f(k) = (\mathcal{F}^{-1}F)(k) = \int_{-1/2}^{1/2} F(\xi)e_k(\xi)d\xi =$ $< F, e_{-k}\chi_{_{I_1}} >_{L^2(\widehat{\mathbb{R}})} = < f, T_k(\mathrm{sinc}) >$. Using $(i)$, $\sum\limits_k f(k)T_k(\mathrm{sinc})$ converges to $f$ in $L^2(\mathbb{R})$. In order to show that we have unconditional pointwise convergence, we fix $x \in \mathbb{R}$ and note that, by unconditional convergence in orthonormal basis expansions of inner products, we have

$$f(x) = < F, e_{-x}\chi_{_{I_1}} > = \sum_{k\in\mathbb{Z}} < F, e_{-k}\chi_{_{I_1}} > \overline{< e_{-x}, e_{-k}\chi_{_{I_1}} >}$$

$$= \sum_{k\in\mathbb{Z}} f(k)\mathrm{sinc}(x-k) = \sum_{k\in\mathbb{Z}} f(k)(T_k\mathrm{sinc})(x) \qquad (2)$$

The corresponding result using values on $\mathbb{Z}+x_0$ in place of values on $\mathbb{Z}$ follows from $(ii)$ by applying (2) to $g = T_{-x_0}f \in \mathbb{B}_1(\mathbb{R})$ and replacing $x$ with $x - x_0$ so $g(k) = f(x_0 + k)$ and $f(x)$ $= g(x - x_0)$, *etc.*

$(iv)$ holds for $z = x \in \mathbb{R}$ since, for $F=\widehat{f}$, $f(x-k)$ is the $(-k)^{th}$ Fourier coefficient of the function $Fe_{-x}\chi_{_{I_1}} \in L^2(I_1)$ and the $L^2$ norm of this function is $\|F\|_{L^2(\widehat{\mathbb{R}})} =$

$||f||_{L^2(\mathbb{R})}$. As we noted above, $f$ has a unique extension to a holomorphic function on $\mathbb{C}$ so $(iv)$ also holds for each $z \in \mathbb{C}$ by analytic continuation.

*Remarks. In the literature, 2.2(iii) is known as the Whittaker-Shannon-Kotel'nikov Sampling Theorem. As the above proof indicates, it's merely an application of Fourier series results. In complex analysis, 2.2(iv) for $f$ = sinc is the basic tool needed to obtain Mittag-Leffler and Weierstrass expansions of various trigonometric functions. There are many other properties of sinc which have proofs analogous to those in $(i)$ – (iv) and are useful in various ways both in real and complex analysis.*

*For signal processing, it's enlightening to compare the properties in Theorem 2.2 for the real analytic Shannon sampling function $\phi_S$=sinc with the properties below for the discontinuous Haar sampling and scaling function* $\phi_H(x) = \chi_{[0,1)}(x)$.

**2.3 Theorem on the Properties of $\phi_H$**

$(i)$ $\{T_k\phi_H : k \in \mathbb{Z}\}$ is an orthonormal basis for the closed subspace $V_{0,H}$ of $L^2(\mathbb{R})$ consisting of pointwise well defined square-integrable functions on $\mathbb{R}$ which are constant on each half-open unit interval $[k, k+1)$ with integer endpoints.

$(ii)$ For $y \in \mathbb{R}$, $T_y$ maps $V_{0,H}$ into itself $\Leftrightarrow y \in \mathbb{Z}$.

$(iii)$ If $S=\{x_k : k \in \mathbb{Z}\}$ with $x_k \in [k, k+1)$ for each $k \in \mathbb{Z}$, then, for each $f \in V_{0,H}$ and each $x \in \mathbb{R}$,

$$f(x) = \sum_{k\in\mathbb{Z}} f(x_k)T_k\phi_H(x) \qquad (3)$$

and the series in (3) also converges to $f$ in $L^2(\mathbb{R})$. Hence, $f$ is uniquely determined by its values on any <u>sample set</u> of the type

S and it is customary to call (3) the sampling equation.

Proof. Unlike the tricky Fourier series arguments needed to prove Theorem 2.2, the proof of Theorem 2.3 follows immediately from the observation that $\mathrm{T}_k\phi_H = \chi_{[k,k+1)}$ is a unit vector in $\mathrm{L}^2(\mathbb{R})$ whose support is disjoint from that of $\mathrm{T}_l\phi_H$ for $k \neq l$ and, for each $x_k \in [k, k+1)$, $f(x_k) = \ < f, \chi_{[k,k+1)} >$, *etc.*

**2.4 Definition.** For $g \in \mathrm{L}^2(\mathbb{R})$,
$\mathrm{G}_g = \{\mathrm{T}_k\mathrm{M}_l\, g : (k, l) \in \mathbb{Z} \times \mathbb{Z}\}$ is the Gabor system with integer translations and modulations generated by $g$.

Note that by Theorem 1.3,
$\mathcal{F}(\mathrm{G}_g) = \{\mathrm{M}_{-k}\mathrm{T}_l\, \widehat{g} = T_l M_{-k}\, \widehat{g} : (l, -k) \in \mathbb{Z} \times \mathbb{Z}\} = \mathrm{G}_{\widehat{g}}$
since $e_l(k) = 1$ for all integers, $k, l$. Since $\mathcal{F}$ is unitary,
$\mathrm{G}_g$ is an orthonormal basis of $\mathrm{L}^2(\mathbb{R}) \Leftrightarrow \mathrm{G}_{\widehat{g}}$ is an orthonormal basis of $\mathrm{L}^2(\widehat{\mathbb{R}})$.

**2.5 Theorem. Both $\mathrm{G}_{\phi_H}$ and $\mathrm{G}_{\phi_S}$ are orthonormal bases for $\mathrm{L}^2(\mathbb{R})$.**

**Proof**. For any unit interval $\mathrm{I} \subset \mathbb{R}$, it's clear from Fourier series considerations that $\{\mathrm{M}_l\chi_I : l \in \mathbb{Z}\}$ is an orthonormal basis of $\mathrm{L}^2(\mathrm{I}) \subset \mathrm{L}^2(\mathbb{R})$. Since the family of translated intervals $I + k$, $k \in \mathbb{Z}$, are, modulo null sets, mutually disjoint with union $\mathbb{R}$, it follows from Theorem 1.3 that $\mathrm{G}_{\chi_I}$ is an orthonormal basis of $\mathrm{L}^2(\mathbb{R})$. Applying this to the intervals $I = [0, 1)$ and $I = I_1 = [-1/2, 1/2)$ and recalling that $\phi_S$ is the Fourier transform of $\chi_{I_1}$, 2.6 follows. Indeed, aside from unimodular scaling factors and identifying $\widehat{\mathbb{R}}$ with $\mathbb{R}$, $\mathrm{G}_{\phi_S}$ can be interpreted as the Fourier transform of $\mathrm{G}_{\phi_H}$.

**2.6 Remarks.**

($i$) One can similarly define higher dimensional Gabor systems $\mathrm{G}_g$ generated by $g \in \mathrm{L}^2(\mathbb{R}^n)$:
$\mathrm{G}_g \equiv \{\mathrm{T}_k\mathrm{M}_l g : k \in \mathbb{Z}^n \subset \mathbb{R}^n,\ l \in (\mathbb{Z}^n)^\perp \subset (\mathbb{R}^n)\hat{}\}$ where
$(\mathbb{Z}^n)^\perp \equiv \{\xi \in (\mathbb{R}^n)\hat{}: \xi k \in \mathbb{Z}$ for each $k \in \mathbb{Z}^n\}$
$= \{$elements in $(\mathbb{R}^n)\hat{}$ with integer entries$\}$.
By $n-$variable Fourier series, if either $g$ or $\widehat{g}$ is the characteristic function of a unit cube, then $\mathrm{G}_g$ is an orthonormal basis for $\mathrm{L}^2(\mathbb{R}^n)$ and $\mathrm{G}_{\widehat{g}}$ is an orthonormal basis for $\mathrm{L}^2((\mathbb{R}^n)\hat{})$.

($ii$) A lattice in $\mathbb{R}^n$ is a subset $\mathcal{L}$ of the form $a\mathbb{Z}^n$ for some $\mathrm{GL}(n,\mathbb{R})$. Thus, $\mathcal{L}$ is the set of linear combinations with integer coefficients of the columns of $a$. The lattice dual $\mathcal{L}^\perp$ of $\mathcal{L}$ is defined to be the set of all $\xi \in (\mathbb{R}^n)\hat{}$ for which $\xi l \in \mathbb{Z}$ for each $l = ak \in \mathcal{L}$. Clearly, $\mathcal{L}^\perp = (\mathbb{Z}^n)^\perp\, a^{-1}$ is the set of linear combinations with integer coefficients of the rows of $a^{-1}$. We can then define $(\mathcal{L},\mathcal{L}^\perp)$ Gabor systems in $\mathrm{L}^2(\mathbb{R}^n)$ by applying to a generator $g$ translations by members of $\mathcal{L}$ and modulations by members of $\mathcal{L}^\perp$; then applying the Fourier transform gives us the $(\mathcal{L}^\perp, \mathcal{L})$ Gabor system in $\mathrm{L}^2((\mathbb{R}^n)\hat{})$ which applies to $\widehat{g}$ translations by members of $\mathcal{L}^\perp$ and modulations by members of $\mathcal{L}$. We get orthonormal bases when either our generator $g$ is the characteristic function of a lattice tiling domain C for $\mathcal{L}$, $i.e.$, $\mathrm{C} \subset \mathbb{R}^n$ is a measurable set for which $\mathbb{R}^n$ is the disjoint union of the "tiles" $\mathrm{C}+l$, $l \in \mathcal{L}$), or $\widehat{g}$ is the characteristic function of a lattice tiling domain $\mathrm{C}'$ for $\mathcal{L}^\perp$. This is not the only way in which Gabor systems can be orthonormal bases, just the easiest way to construct orthonormal Gabor systems. With $[0,1)^n$ turned into a subset of $\mathbb{R}^n$, $a[0,1)^n$ is an easy example of an $a\mathbb{Z}^n-$tiling domain; similarly, with $[0,1)^n$ turned into a subset of

$(\mathbb{R}^n)\hat{}$, $[0,1)^n a^{-1}$ is an example of an $(a\mathbb{Z}^n)^{\perp}$ – tiling domain $\mathbb{R}^n$. Sadly, one can show there is no orthonormal Gabor system whose generator is a smooth, compactly supported function. This is a very big drawback for applications of Gabor systems to signal analysis; Gabor systems are simply not efficient in the sense that, for most signals, we need a large number of coefficients in order to be able to reconstruct a reasonably good approximation of our signal.

*(iii)* There are also higher dimensional versions of Haar and Shannon functions with orthonormal basis and sampling properties analogous to those in 2.2 and 2.3 We won't take the time to belabor this.

2.7 **Definitions.**

*(i)* The basic dyadic dilation operators on $L^2(\mathbb{R})$ are $D=D_{1/2}$ and $D^{-1} = D_2$ with $\{D^j : j \in \mathbb{Z}\}$ the group of dyadic dilation operators generated by D.

*(ii)* For $\psi \in L^2(\mathbb{R})$,
$WAV_\psi \equiv \{\psi_{j,k} = D^j T_k \psi : (j,k) \in \mathbb{Z} \times \mathbb{Z}\}$ is the <u>wavelet system with integer translations and dyadic dilations</u> generated by $\psi$. If $WAV_\psi$ is an orthonormal basis for $L^2(\mathbb{R})$, we say $\psi$ is a <u>dyadic orthonormal wavelet.</u>

*(iii)* A <u>dyadic, orthonormal, multi-resolution analysis</u> (for short, a dyadic ON MRA) for $L^2(\mathbb{R})$ is a family $(V_j)_{j\in\mathbb{Z}}$ of closed subspaces of $L^2(\mathbb{R})$ for which the following properties hold:

MRA (1) (nested property) $V_j \subset V_{j+1}\, \forall j \in \mathbb{Z}$;
MRA (2) (intersection property) $\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$;
MRA (3) (union property) $\bigcup_{j\in\mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$;
MRA (4) (dyadic dilation property) $V_j = D^j V_0\, \forall j \in \mathbb{Z}$;

MRA (5) (orthonormal scaling property) There is a function $\phi \in \mathrm{V}_0$ for which $\{\mathrm{T}_k\phi : k \in \mathbb{Z}\}$ is an orthonormal basis for $V_0$.

There are also non-ON dyadic MRAs where we still have MRA(1) – MRA(4) but replace MRA(5) with a weaker condition. In higher dimensions, with $\mathbb{R}$ and $\mathbb{Z}$ replaced by $\mathbb{R}^n$ and $\mathbb{Z}^n$, we have to make a choice of $a \in \mathrm{GL}(n, \mathbb{R})$ for which MRA (4) holds for $\mathrm{D}=\mathrm{D}_{a^{-1}}$. For best results, attention is limited to $a's$ with integer entries and with all of their eigenvalues having magnitude strictly greater than 1.

.

2.8. **Remarks.** When $(\mathrm{V}_j)_{j\in\mathbb{Z}}$ is a dyadic MRA, MRA(1)-(4) imply that, for $\mathrm{W}_0 \subset \mathrm{V}_1$ the orthogonal complement of $\mathrm{V}_0 \subset \mathrm{V}_1$, $\mathrm{W}_j = \mathrm{D}^j\mathrm{W}_0$ is the orthogonal complement of $\mathrm{V}_j$ in $\mathrm{V}_{j+1}$ and $\mathrm{L}^2(\mathbb{R})$ is the orthogonal direct sum of the family of the closed subspaces $\mathrm{W}_j, j \in \mathbb{Z}$. In the book *The Mathematical Theory of Wavelets* by E. Hernandez and G. Weiss, it's shown that we can always select $\psi \in W_0$ for which $\{\mathrm{T}_k\psi : k \in \mathbb{Z}\}$ is an orthonormal basis for $\mathrm{W}_0$. Later, we'll go over the details of this and also will show that $\psi$ is unique up to replacement of $\widehat{\psi}$ by $\mu\widehat{\psi}$ for $\mu$ a unimodular, $\mathbb{Z}$-periodic function. It follows that $\mathrm{WAV}_\psi$ is an orthonormal basis for $\mathrm{L}^2(\mathbb{R})$ for each such $\psi$. In this way, dyadic ON MRA's give rise to dyadic orthonormal wavelets which are uniquely determined by the MRA modulo unimodular $\mathbb{Z}$-periodic Fourier domain multipliers. Theorem 2.9 illustrates this general result.

**2.9 Theorem on the Haar and Shannon wavelets.**

($i$) The Haar wavelet function $\psi_H \equiv \chi_{[0,1/2)} - \chi_{[1/2,1)}$ is an ON dyadic MRA wavelet with scaling function $\phi_H$.

($ii$) The Shannon wavelet function $\psi_S = \mathcal{F}^{-1}(\chi_{[-1,-1/2)\cup[1/2,1)})$

is an ON dyadic MRA wavelet with scaling function $\phi_S$.

**Proof.** $(i)$ As we saw in Theorem 2.3, $V_{0,H}$ consists of the functions in $L^2(\mathbb{R})$ which are constant on half-open intervals with integer endpoints and has $\{T_k\phi_H : k \in \mathbb{Z}\}$ as an orthonormal basis. Define $V_{j,H}$ to be $D^j(V_{0,H})$ for $j \in \mathbb{Z}$. The members of $V_j$ are the functions in $L^2(\mathbb{R})$ which are constant on intervals of the form $I_{j,k} = [k/2^j, (k+1)/2^j]$. For $j = 1$, each such interval is contained in an interval with integer endpoints so $V_{0,H} \subset V_{1,H}$ and it follows that $V_{j,H} \subset V_{j+1,H}$ for all $j \in \mathbb{Z}$. This checks the MRA conditions (1), (4), and (5) for $(V_{j,H})_{j\in\mathbb{Z}}$. For $j < 0$, members of $V_{j,H}$ are constant on the half-open interval $[2^{|j|-1}, 2^{|j|+1})$ centered at 0 and having length $2^{|j|}$. As $j \to -\infty$, these intervals exhaust $\mathbb{R}$ and it follows that MRA(2) is satisfied. Finally, MRA (3) follows by the ability to get increasingly good approximations to any function in $L^2(\mathbb{R})$ by simple functions in $V_{j,H}$ as $j \to \infty$. Clearly $\psi_H$ is in $V_{1,H}$ and is perpendicular to $T_k\phi_H$ for each $k \in \mathbb{Z}$, so, in the notations of 2.9, $\psi_H \in W_{0,H}$. Easy direct computations show that $\{T_k\psi_H : k \in \mathbb{Z}\}$ is an orthonormal basis of $W_{0,H}$ so $\psi_H$ is a member of the family of orthonormal, dyadic wavelets determined by the Haar MRA system $(V_{j,H})_{j\in\mathbb{Z}}$.

$(ii)$ Here we must do all of our checking in the Fourier domain, remembering that $D = D_{1/2}$ implies $\widehat{D} = D_2$. Hence, for $j > 0$, application of $\widehat{D}^j$ dilates supports by $2^j$, while, as in $(i)$, application of $D^j$ contracts supports by $\frac{1}{2^j}$. By Theorem 2.3, $V_{0,S}$ $= \mathbb{B}_1(\mathbb{R})$ has $\{T_k\phi_S : k \in S\}$ as an orthonormal basis and $(V_{0,S})^\wedge = \mathcal{F}(V_{0,S}) = L^2(I_1) \subset L^2(\widehat{\mathbb{R}})$. Let $I_j = 2^j I_1 = [-2^{j-1}, 2^{j-1})$. We then have $(V_{j,S})^\wedge = \widehat{D}^j(V_{0,S})^\wedge$ $= L^2(I_j)$ and we readily deduce from this that $(V_{j,S})_{j\in\mathbb{Z}}$ satisfies

MRA(1)-(5). Furthermore, it's obvious from Fourier series considerations that $(W_{0,S})^\frown = L^2([-1, -1/2) \cup [1/2, 1))$ has $\{e_{-k}(\psi_S)^\frown : k \in \mathbb{Z}\}$ as an orthonormal basis so $W_{0,S}$ has $\{T_k\psi_S : k \in \mathbb{Z}\}$ as an orthonormal basis and hence $\psi_H$ is in the family of orthonormal, dyadic wavelets defined by the Shannon MRA $(V_j, S)_{j\in\mathbb{Z}}$.

**2.10 MRA Wavelet Applications.** When $(V_j)_{j\in\mathbb{Z}}$ is a dyadic MRA with scaling function $\phi$ and wavelet function $\psi$, we have, for each $j > 0$, $V_j = V_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_{j-1}$. For $f = f_j \in V_j$, let $f_0$ be the component of $f$ in $V_0$ and $E_i, 0 \le i \le j-1$, the component of $f$ in $W_i$. Then $f_0$ is the orthogonal projection of $f$ to $V_0$ and, for $1 \le i \le j-1$, $f_i = f_0 + E_0 + \cdots + E_{i-1}$ is the orthogonal projection of $f$ to $V_i$. We think of $f_i$ as the $i^{th}$ resolution level approximation of $f$ and think of $E_i = f_{i+1} - f_i$ as the error term in passing from $f_{i+1}$ down to $f_i$.

For the Shannon MRA, we pass from $f_i$ to $f_{i-1}$ by "filtering out" the frequencies in $J_i = I_i \backslash I_{i-1}$; thus $(f_{i-1})^\frown = \chi_{J_i}\widehat{f_i}$. For the Haar MRA, each of the intervals $I_{i-1,k} = [k/2^{i-1}, (k+1)/2^{i-1})$ is the disjoint union of the intervals $I_{i,2k}$ and $I_{i,2k+1}$ so, for each $k$, the constant value of $f_{i-1}$ on $I_{i-1,k}$ is the average of the constant values of $f_i$ on $I_{i,2k}$ and $I_{i,2k+1}$. This makes it very easy to pass from $f_i$ to $f_{i-1}$ and to calculate $E_{i-1} = f_i - f_{i-1}$ for both the Haar and Shannon MRAs.

In practical applications, the only information we have about a one-dimensional signal $f$ is a finite collection of data values. We then pick a convenient dyadic ON MRA $(V_j)_{j\in\mathbb{Z}}$ with scaling function $\phi$ and associated wavelet function $\psi$. For a reasonably large $j > 1$, we assume our signal is in $V_j$ and turn our data values into the coefficients for $f$ in terms of the orthonormal basis $\{D^jT_k\phi : k \in \mathbb{Z}\}$ for $V_j$. As we'll discuss below, there is a very fast way to calculate the coefficients of each $E_i$ relative to the orthonormal basis $\{D^iT_k\psi : k \in \mathbb{Z}\}$ for $W_i$ and to calculate the coefficients of $f_0$ relative to $\{T_k\phi : k \in \mathbb{Z}\}$; in brief, we never have to compute inner products to find these new coefficients but instead can use algebraic manipulations on our initial collection of coefficients. We can then store all of these coefficients in our "home" computer and, when asked by someone to send information from which a low level approximation to $f$ can be reassembled, we transmit only the coefficients needed to reproduce $f_i$ for a fairly small value of $i$ : only rarely do we need to pull out and transmit the error coefficients needed to correct $f_i$ to $f_{i'}$ for a "moderately" large value of $i'$. In practice, we almost never need to go back to the data used to obtain $V_j$ coefficients of our original $f$. Also, given a signal for which we have data values, we can go to a library of MRA systems to select a system well adapted to our signal, *i.e.*, one for which it seems reasonable to regard our data values as giving us coefficients for a member of $V_j$ for some reasonable large $j$. This brief explanation of efficiency in coding (small number of coefficients needed for low level resolution approximations) and flexibility (ability to use systems well adapted to the signals of interest) is why wavelet systems have been adopted by engineering journals and engineering associations as the industrial standard for data analysis and all earlier systems are considered outmoded and unacceptable for use in publications. In particular, although Gabor systems are mathematically appealing, they are of little practical use for data analysis since they're not efficient and don't have the features just discussed for wavelet systems. The Haar and Shannon wavelet systems are relatively easy to describe but there are other wavelet systems with much better features for certain applications. In particular, Y. Meyer constructed a large family of dyadic ON MRA wavelets whose Fourier transforms are smooth and, like the Shannon wavelet, have compact support. I. Daubechies constructed, for each $k \in \mathbb{N}$, a dyadic ON MRA wavelet having $k$ continuous derivatives and, like the Haar wavelet, having compact support in the time domain. The Daubechies wavelets are the wavelets now most commonly used in applications. Because of a large host of theorems, one-dimensional mathematical wavelet theory is now "nearly" complete. However, there remain many open questions in higher-dimensional wavelet theory; this is a very active research area.

## 3. Shift Invariant Spaces

**3.1 Overview.** We saw in Section **2** that the starting point for construction of Gabor and dyadic wavelet systems on $L^2(\mathbb{R})$ is the set of translates $T_k\phi$, $k \in \mathbb{Z}$ for some generator $\phi$. When we apply $D^j$ to each of these translates and use the Theorem 1.4 transformation formulas,we are looking at the translates $T_{k/2^j}(D^j\phi)$ of the dilation $D^j\phi$ of $\phi$; in effect we replace the lattice $\mathbb{Z}$ of $\mathbb{R}$ with the lattice $\frac{1}{2^j}\mathbb{Z}$ whenever we apply $D^j$. Similarly, in higher idimensions, we could start off looking at the set of translates of some generating function $\phi \in L^2(\mathbb{R}^n)$ by members of $\mathbb{Z}^n$ but then when we apply $D_a$ and use $D_a T_k \psi = T_{ak} D_a \psi$, we are translating the new generator $D_a\phi$ by members of the lattice $a\mathbb{Z}^n$. In principle, we could use a change of coordinates in which $a\mathbb{Z}^n$ is expressed by $\mathbb{Z}^n$ in the new coordinate system. It's much better not to keep changing coordinates but, instead, to use coordinate-free techniques.

### 3.2 Definitions

$(i)$ When $\mathcal{L}$ is a lattice in $\mathbb{R}^n$, an $\mathcal{L}$ <u>shift-invariant space</u> is a closed subspace V of $L^2(\mathbb{R}^n)$ for which $T_l(V)=V$ for each $l \in \mathcal{L}$ and V is then said to be a <u>principal</u> (or <u>cyclic</u>) $\mathcal{L}$ <u>shift-invariant space</u> if there is some non-zero $\phi \in V$ for which $V = <\phi>_{\mathcal{L}} \equiv$ smallest closed subspace of $L^2(\mathbb{R}^n)$ containing $\{T_l\phi : l \in \mathcal{L}\}$. Then each such $\phi$ is an $\mathcal{L}$ <u>shift-invariant space generator</u> for V.

$(ii)$ When V $=$ $< \phi >_{\mathcal{L}}$ is a principal $\mathcal{L}$ shift-invariant space, $\widehat{\mathrm{V}} = \mathcal{F}(V)$ is the image in $\mathrm{L}^2((\mathbb{R}^n)\hat{})$ of V under the Fourier transform $\mathcal{F}$. Since $(\mathrm{T}_l\phi)\hat{} = e_{-l}\widehat{\phi}$, $\widehat{\mathrm{V}}$ is the smallest closed subspace of $\mathrm{L}^2((\mathbb{R}^n)\hat{})$ containing $\{e_l\widehat{\phi} : l \in \mathcal{L}\}$. Hence $\{m\widehat{\phi} : m$ is a finite linear combintion of the elementary $\mathcal{L}^{\perp}$ – periodic functions $e_l, l \in \mathcal{L}\}$ is a dense linear subspace of $\widehat{\mathrm{V}}$.

$(iii)$ For $\mathcal{L}$ a lattice in $\mathbb{R}^n$ and $\phi$,$\psi$ in $\mathrm{L}^2(\mathbb{R}^n)$, the $\mathcal{L}$ bracket $[\phi,\psi]_{\mathcal{L}}$ is the $\mathcal{L}^{\perp}$ periodization of $\widehat{\phi}\,\overline{\widehat{\psi}}$, $i.e.$ $[\phi,\psi]_{\mathcal{L}}$ is the $\mathcal{L}^{\perp}$-periodic function on $(\mathbb{R}^n)\hat{}$ which is $a.e.$ well defined by

$$[\phi,\psi]_{\mathcal{L}}(\xi) = \sum_{j\in \mathcal{L}^{\perp}} (\,\widehat{\phi}\,\overline{\widehat{\psi}}\,)(\xi + j). \tag{4}$$

Note that $(\phi,\psi) \mapsto [\phi,\psi]_{\mathcal{L}}$ is $\mathbb{C}$-linear in $\phi$, conjugate $\mathbb{C}$-linear in $\psi$, $[\psi,\phi]_{\mathcal{L}}=\overline{[\phi,\psi]_{\mathcal{L}}}$ , and $[\phi,\phi]_{\mathcal{L}} \geq 0$ $a.e.$ with $[\phi,\phi]_{\mathcal{L}}=0$ $a.e.$ $\Leftrightarrow$ $\phi$ is the zero element in $\mathrm{L}^2(\mathbb{R}^n)$. Aside from the fact that $[\phi,\psi]_{\mathcal{L}}$ is a function rather than a complex number, these are the properties of an inner product. By applying the Cauchy-Schwartz inequality for $l^2(\mathcal{L}^{\perp})$, we also have a Cauchy-Schwartz inequality for brackets:

$$|[\phi,\psi]_{\mathcal{L}}| \leq [\phi,\phi]_{\mathcal{L}}^{\frac{1}{2}}\,[\psi,\psi]_{\mathcal{L}}^{\frac{1}{2}} \quad a.e. \tag{5}$$

As we'll see, $[\,\cdot\,,\cdot\,]_{\mathcal{L}}$ has many other properties analogous to those of inner products.

**3.3 Elemenary Properties of $\mathcal{L}$ brackets and principal $\mathcal{L}$-shift invariant spaces.**

Fix a lattice $\mathcal{L}$ in $\mathbb{R}^n$ and let $\mathbb{T}_{\mathcal{L}^{\perp}}$ be the compact Abelian group $(\mathbb{R}^n)\hat{}\,/\mathcal{L}^{\perp}$. As a set, we can identify $\mathbb{T}_{\mathcal{L}^{\perp}}$ with any $\mathcal{L}^{\perp}$ tiling domain C and we can regard integrable functions on $\mathbb{T}_{\mathcal{L}^{\perp}}$ as locally integrable , $\mathcal{L}^{\perp}$-periodic functions F on $(\mathbb{R}^n)\hat{}$. We denote by

$\int_{\mathbb{T}_{\mathcal{L}^\perp}}$ F d$\xi$ the common value of $\int_C F(\xi)d\xi$ for C any $\mathcal{L}^\perp$ tiling domain. From the theory of Fourier series, we also know that $\{e_k : k \in \mathcal{L}\}$ is an orthonormal basis for $L^2(\mathbb{T}_{\mathcal{L}^\perp})$.

**(*i*) For $\phi$,$\psi$ in $L^2(\mathbb{R}^n)$, $[\phi,\psi]_\mathcal{L}$ is the unique member of $L^1(\mathbb{T}_{\mathcal{L}^\perp})$ whose $k^{th}$ Fourier coefficient $\int_{\mathbb{T}_{\mathcal{L}^\perp}} e_{-k}[\phi,\psi]_\mathcal{L} d\xi$ is $<T_k\phi, \psi>$ .**

**Proof.** By the Plancherel theorem for $\mathbb{R}^n$, $\widehat{\phi}\,\overline{\widehat{\psi}}$ is an integrable function on $(\mathbb{R}^n)\hat{}$ and its integral over $(\mathbb{R}^n)\hat{}$ is equal to $<\phi, \psi>$ . Choose an $\mathcal{L}^\perp$ tiling domain C. Using (4), the translation invariant property of Lebesgue measure, and the Fubini Theorem to interchange $\int$ and $\sum$, we obtain

$$\int_{(\mathbb{R}^n)\hat{}} \widehat{\phi}\,\overline{\widehat{\psi}}\,(\xi)d\xi = \sum_{j\in\mathcal{L}^\perp} \int_C (\widehat{\phi}\,\overline{\widehat{\psi}})(\xi + j)d\xi = \int_C [\phi,\psi]_\mathcal{L}(\xi)d\xi. \quad (6)$$

Using (5), (6), and the Cauchy-Schwartz inequality for $L^2\left(\mathbb{T}_{\mathcal{L}^\perp}\right)$, it follows that $[\phi,\psi]_\mathcal{L}$ is in $L^1(\mathbb{T}_{\mathcal{L}^\perp})$ with

$$\int_{\mathbb{T}_{\mathcal{L}^\perp}} [\phi,\psi]_\mathcal{L}\, d\xi = \int_{(\mathbb{R}^n)\hat{}} \widehat{\phi}\,\overline{\widehat{\psi}}\,(\xi)d\xi = <\phi, \psi> \quad (7)$$

and $\| [\phi,\psi]_\mathcal{L}\|_{L^1(\mathbb{T}_{\mathcal{L}^\perp})} \le ||\phi||_{L^2(\mathbb{R}^n} ||\psi||_{L^2(\mathbb{R}^n}$.

We now recall that $(T_k\phi)\hat{} = e_{-k}\phi\hat{}$. Since $e_{-k}$ is $\mathcal{L}^\perp$-periodic, we obtain from (4) that

$$[T_k\phi,\psi]_\mathcal{L} = e_{-k}\,[\phi,\psi]_\mathcal{L} \quad (8)$$

so substituting $T_k\phi$ for $\phi$ in (6) completes the proof of (*i*).

**(*ii*) $\mathcal{B}_{\mathcal{L},\phi} = \{T_k\phi : k \in \mathcal{L}\}$ is an orthonormal basis of $<\phi>_\mathcal{L}$ $\Leftrightarrow$ $[\phi,\phi]_\mathcal{L} = 1$ *a.e.***

**Proof.** Using (*i*), $[\phi,\phi]_\mathcal{L} = 1$ *a.e.* $\Leftrightarrow$ its $0^{th}$ Fourier coefficient is 1 and all of its other Fourier coefficients are 0

$\Leftrightarrow \forall k, l \in \mathcal{L},\ < \mathrm{T}_k\phi, \mathrm{T}_l\phi > \ = <\mathrm{T}_{k-l}\phi, \phi >$ is 1 when $k = l$ and is 0 when $k \neq l$

$\Leftrightarrow \mathcal{B}_{\mathcal{L},\phi}$ is an orthonormal basis for the closure $<\phi>_{\mathcal{L}}$ of its span.

**(*iii*) For $\phi$,$\psi$ in $\mathrm{L}^2(\mathbb{R}^n)$, $[\phi,\psi]_{\mathcal{L}} = 0$ a.e. $\Leftrightarrow <\phi>_{\mathcal{L}}$ and $<\psi>_{\mathcal{L}}$ are perpendicular subspaces of $\mathrm{L}^2(\mathbb{R}^n)$.**

**Proof.** This follows from $(i)$ since $<\phi>_{\mathcal{L}}$ and $<\psi>_{\mathcal{L}}$ are perpendicular subspaces

$\Leftrightarrow 0 = < \mathrm{T}_k\phi, \mathrm{T}_l\psi > \ = <\mathrm{T}_{k-l}\phi, \psi > \forall k, l \in \mathcal{L}$.

**(*iv*) When $\phi$ is a generator for a principal $\mathcal{L}$ shift-invariant space V, then $\widehat{\mathrm{V}} = \{m\widehat{\phi} : m$ is measurable and $\mathcal{L}^{\perp}$-periodic with $|m|^2[\phi,\phi]_{\mathcal{L}} \in \mathrm{L}^1(\mathbb{T}_{\mathcal{L}^{\perp}})\}$.**

**Proof.** When $m\widehat{\phi}$ is in the set described above, the $\mathcal{L}^{\perp}$-periodicity of $m$ gives $\sum_{j\in\mathcal{L}^{\perp}} |m(\xi + j)\widehat{\phi}(\xi + j)|^2 = |m(\xi)|^2$ $[\phi,\phi]_{\mathcal{L}}(\xi)$ $a.e.$ and hence the same integral periodization as in $(i)$ implies that $\widehat{f} = m\widehat{\phi}$ is in $\mathrm{L}^2((\mathbb{R}^n)^{\widehat{}}\ )$ so $\widehat{f}$ is the Fourier transform of some $f \in \mathrm{L}^2(\mathbb{R}^n)$ with

$$\|f\|^2_{L^2(\mathbb{R}^n)} = \|\widehat{f}\|_{L^2((\mathbb{R}^n)^{\widehat{}})} = \int_{\mathbb{T}_{\mathcal{L}^{\perp}}} |m|^2[\phi,\phi]_{\mathcal{L}} d\xi \qquad (9)$$

It's obvious that the collection of these functions $\widehat{f} = m\widehat{\phi}$ is a closed subspace of $\mathrm{L}^2((\mathbb{R}^n)^{\widehat{}}\ )$ and is the closure of the dense linear subspace of $\widehat{\mathrm{V}}$ described in 3.2$(ii)$.

**(*v*) Each principal $\mathcal{L}$ shift-invariant space V uniquely determines, modulo a set of measure zero, a measurable subset $\Omega$=$\Omega_{\mathrm{V}}$ of $\mathbb{T}_{\mathcal{L}^{\perp}}$ such that, for any $\psi \in$ V, $[\psi,\psi]_{\mathcal{L}}$ is 0 $a.e.$ off $\Omega$ and $\psi$ generates V $\Leftrightarrow [\psi,\psi]_{\mathcal{L}}>0$ $a.e.$ on $\Omega$. Furthermore, we can always choose generators $\psi$ for V with $[\psi,\psi]_{\mathcal{L}}=\chi_{\Omega}\, a.e.$ and then $\mathcal{B}_{\mathcal{L},\psi}$ is a <u>Parseval frame</u> for V in the sense that**

$$\sum_{j\in\mathcal{L}} |< f, \mathrm{T}_k\psi > |^2 = ||f||^2_{\mathrm{L}^2(\mathbb{R}^n)} \, \forall f \in V. \tag{10}$$

**Proof.** Choose any generator $\phi$ for V and define $\Omega$ to be $supp\,[\phi,\phi]_{\mathcal{L}} = \{\xi + \mathcal{L}^\perp : [\phi,\phi]_{\mathcal{L}}(\xi) > 0\}$. For $\psi \in$ V, $(iv)$ tells us that $\widehat{\psi}$=$m\widehat{\phi}$ with $|m|^2\,[\phi,\phi]_{\mathcal{L}}$ integrable on $\mathbb{T}_{\mathcal{L}^\perp}$. Using (4) and periodicity of $m$, $[\psi,\psi]_{\mathcal{L}} = |m|^2\,[\phi,\phi]_{\mathcal{L}}$ so $[\psi,\psi]_{\mathcal{L}} = 0$ $a.e.$ off $\Omega$. When $[\psi,\psi]_{\mathcal{L}}$>0 $a.e.$ on $\Omega$, we have $m$ non-zero $a.e.$ $\Omega$; for $f \in V$, we can then use $(iv)$ to describe $\widehat{f}$ either in the form $m'\widehat{\phi}$ or the form $m''\widehat{\psi}$ with $m'' = \frac{m'}{m}\chi_\Omega$ and to deduce that $\psi$ generates V.
When we define $\psi$ to be the member of V for which $\widehat{\psi}$=$\frac{\chi_\Omega}{[\phi,\phi]_{\mathcal{L}}{}^{1/2}}\widehat{\phi}$, we have $[\psi,\psi]_{\mathcal{L}}$=$\chi_\Omega a.e.$ Using (i), for any $f \in V$, $\widehat{f} = m\widehat{\psi}$ with $m \in L^2(\Omega) \equiv \{$members of $\mathrm{L}^2(\mathbb{T}_{\mathcal{L}^\perp})$ vanishing $a.e.$ off $\Omega\}$. By $(i)$ and the Plancherel Formula for Fourier series, $||f||^2 = ||\widehat{f}||^2 = ||m||^2_{\mathrm{L}^2(\mathbb{T}_{\mathcal{L}^\perp})}$ = sum of the squares of the Fourier coefficients for $m = \sum_{k\in\mathcal{L}} |< f, \mathrm{T}_k\psi > |^2$. Hence, $\mathcal{B}_{\mathcal{L},\psi}$ is a Parseval frame for V.

*Note: Given V as in $(v)$, $(ii)$ implies that there exists a generator $\psi$ for V with $\mathcal{B}_{\mathcal{L},\psi}$ an ON basis of V $\Leftrightarrow$ $\Omega_V = \mathbb{T}_{\mathcal{L}^\perp}$ modulo a null set. In general, the best we can do is use a Parseval frame generator $\psi$ for V. This is still very nice since, among other properties, for each $f \in V$, we have $\widehat{f} = [f,\psi]_{\mathcal{L}}\widehat{\psi}$ and it follows that, for $f \in L^2(\mathbb{R}^n)$, the orthogonal projection of $f$ onto V is the unique function $g$ for which $\widehat{g} = [f,\psi]_{\mathcal{L}}\widehat{\psi}$.*

**Definition.** When V is an arbitrary $\mathcal{L}$ shift-invariant subspace of

$L^2(\mathbb{R}^n)$, a countable subset $\Psi=\{\phi_i : i \in I\}$ is an orthogonal Parseval frame generating set (OPFGS) for V if the following hold:

$(i)$ $V=\bigoplus_{i\in I} < \psi_i >_{\mathcal{L}}$ (orthogonal direct sum);

$(ii)$ For each $i \in I$, $\mathcal{B}_{\mathcal{L},\psi_i}$ is a Parseval frame for $< \psi_i >_{\mathcal{L}}$ and, hence, by $(i)$, $\mathcal{B}_{\mathcal{L},\Psi} \equiv \bigcup_{i\in I}\mathcal{B}_{\mathcal{L},\psi_i}$ is a Parseval frame for V. Noting that the orthogonal complement in V of any $\mathcal{L}$ shift-invariant subspace $W \subset V$ is also a $\mathcal{L}$ shift-invariant subspace (because $T_k$ is a unitary operator on V for each $k \in \mathcal{L}$), the existence of OPFGSs for V follows by Zorn's Lemma. Alternatively, one can use a bracket version of the Gram-Schmidt process to convert any countable generating set for V to an OPFGS.

We denote by $r_{V,\mathcal{L}}$ the smallest cardinal number for which V can be described as the orthogonal direct sum of $r_{V,\mathcal{L}}$ principal $\mathcal{L}$ shift-invariant spaces. Borrowing some terminology from abstract algebra, we call $r_{V,\mathcal{L}}$ the rank of V over $\mathcal{L}$.

**3.4 Dimension Theorem for Shift Invariant Spaces. Fix $\mathcal{L}$ as in 3.5 and let V be any $\mathcal{L}$ shift-invariant subspace of $L^2(\mathbb{R}^n)$. Then there exists a measurable function $dim_{V,\mathcal{L}}:\mathbb{T}_{\mathcal{L}^\perp}\to\mathbb{N} \cup \{0,\infty\}$ with the following properties:**

**$(i)$ For any OPFGS $\Psi=\{\psi_i : i \in I\}$ for V,**

$$dim_{V,\mathcal{L}} = \sum_{i\in I}[\psi_i, \psi_i]_{\mathcal{L}} \ a.e. \ ; \qquad (11)$$

**$(ii)$ $\|dim_{V,\mathcal{L}}\|_\infty = r_{V,\mathcal{L}}$;**

**$(iii)$ When $r_{V,\mathcal{L}}=N \in \mathbb{N}$ and, for $1 \leq j \leq N$, $\Omega_j = \{\xi \in \mathbb{T}_{\mathcal{L}^\perp} : dim_{V,\mathcal{L}}(\xi) \geq j\}$, we can choose an OPFGS $\Psi=\{\psi_j : 1 \leq j \leq N\}$ for V such that $[\psi_j, \psi_j]_{\mathcal{L}}=\chi_{\Omega_j}$ for each $j$.**

**Proof.** *(i)* We merely need to imitate the easy proof using inner products of the dimension theorem for Hilbert spaces. Thus, suppose $\Phi=\{\phi_i : i \in I\}$ and $\Psi=\{\psi_j : j \in J\}$ are OPFGSs for V. From the Note after 3.3(v), we have another parallel between brackets and inner products, namely, for each $f \in V$,

$$[f,f]_{\mathcal{L}} = \sum_{i\in I}|[f,\phi_i]_{\mathcal{L}}|^2 = \sum_{j\in J}|[f,\psi_j]_{\mathcal{L}}|^2. \tag{12}$$

We can then apply (12) to each member of $\Phi$ $(respectively,\ \Psi)$ using brackets with members of $\Psi$ $(respectively,\ \Phi)$ and add up the results in either order with the aid of the Hermitian symmetry property for brackets :

$$\sum_{i\in I}[\phi_i,\phi_i]_{\mathcal{L}} = \sum_{i\in I}\sum_{j\in J}|[\phi_i,\psi_j]_{\mathcal{L}}|^2$$
$$= \sum_{j\in J}\sum_{i\in I}|[\psi_j,\phi_i]_{\mathcal{L}}|^2 = \sum_{j\in J}[\psi_j,\psi_j]_{\mathcal{L}}.$$

Clearly, $(i)$ follows.

**Sketch of the Proof of *(ii)* and *(iii)*.** $(iii)$ follows from a somewhat lengthy "rearrangement" process based on checking that $<\psi>_{\mathcal{L}} + <\psi'>_{\mathcal{L}} = <\psi+\psi'>_{\mathcal{L}}$ when $(supp\,[\psi,\psi]_{\mathcal{L}}) \cap (supp\,[\psi',\psi']_{\mathcal{L}})$ is a null set. There is actually a version of $(iii)$ when $r_{V,\mathcal{L}}=\infty$. Then $(ii)$ follows easily from $(iii)$.

**3.5 Corollary. Let V be as in 3.4.**

**$(i)$ If $r_{V,\mathcal{L}}$ is finite, there exists an OPFSG $\Psi$ for V such that $\mathcal{B}_{\Psi,\mathcal{L}}$ is an orthonormal basis of B $\Leftrightarrow dim_{V,\mathcal{L}} = r_{V,\mathcal{L}}$ $a.e.$ and then the OPFGSs with this orthonormal property are precisely those with $r_{V,\mathcal{L}}$ members.**

**$(ii)$ When $V', V''$ are $\mathcal{L}$ shift-invariant subspaces of V for which $V=V' \oplus V''$ then $dim_{V,\mathcal{L}}=dim_{V',\mathcal{L}} + dim_{V'',\mathcal{L}}$ $a.e.$**

**Proof.** Both $(i)$ and $(ii)$ are immediate from Theorem 3.4.

### 3.6 Zak Transforms and $\mathrm{L}^2$ – Sampling Functions

**$(i)$ Definition.** Relative to the lattice $\mathbb{Z}^n$, the Zak Transform $\mathrm{Z}f$ of $f \in \mathrm{L}^2(\mathbb{R}^n)$ is the $a.e.$ well defined function from $\mathbb{R}^n$ into $\mathrm{L}^2(\mathbb{T}_{(\mathbb{Z}^n)^\perp})$ described by the Fourier series

$$(\mathrm{Z}f)(x,\,\cdot\,) = \sum_{k\in\mathbb{Z}^n} f(x+k)e_{-k}(\,\cdot\,) \tag{13}$$

Thus, as we used above for $\widehat{f}$, $((f(x+k))_{k\in\mathbb{Z}^n}$ is a square summable sequence for $a.e.\ x$. By an easy change-of-summation-index computation, $(\mathrm{Z}f)(x+l,\ \cdot\,) = e_l(\,\cdot\,)(\mathrm{Z}f)(x,\ \cdot\,)\ a.e.$ so $|\mathrm{Z}f|$ is $\mathbb{Z}^n \times (\mathbb{Z}^n)^\perp$ periodic. A routine computation shows that for C any $\mathbb{Z}^n$ tiling domain in $\mathbb{R}^n$ and C$'$ any $(\mathbb{Z}^n)^\perp$ tiling domain in $(\mathbb{R}^n)\widehat{}$,

$$\int_{\mathbb{R}^n}|f|^2dx = \int_C\int_{C'}|\mathrm{Z}f|^2d\xi dx \tag{14}$$

One can then show that $f \mapsto \mathrm{Z}f$ is unitary from $\mathrm{L}^2(\mathbb{R}^n)$ onto the Hilbert space of functions satisfying the above transformation condition and having magnitudes which are square integrable on C $\times$ C$'$. It's clear from (13) that Z can be interpreted as a discretized Fourier transform. In fact, there is an elementary proof of the Plancherel Theorem for $\mathbb{R}^n$ based on interpreting $(\mathcal{F}f)(\xi)$ as the average over $x$ of the quantitities $e^{-2\pi i\xi x}(\mathrm{Z}f)(x,\xi)$. For many purposes, Zak transforms "accomplish" the same things as Fourier transforms but are much easier to compute and to invert. As an example of this,

$$[f,g]_{\mathbb{Z}^n}(\xi) = \int_C(\mathrm{Z}f)(x,\xi)\overline{(\mathrm{Z}g)(x,\xi)}\,dx \tag{15}$$

in view of the uniqueness characterization of brackets in 3.2 and an easy calculation showing that $< \mathrm{T}_k f, g >$ is obtained by multiplying the right side of (15) by $e_{-k}(\xi)$ and integrating over $\mathrm{C}'$. There are many other uses for Zak transforms in the theory of harmonic analysis and in applied harmonic analysis. In $(iv)$, we will mention one of these applications.

$(\boldsymbol{iii})$ We observed in Section 2 that the Haar and Shannon scaling functions $\phi_H$ and $\phi_S$ are also sampling functions in the sense that members of $\mathrm{V}_{0,H}$ and $\mathrm{V}_{0,S}$ are uniquely determined by their values on $\mathbb{Z}$ and these values can be used as coefficients in expressing functions in these spaces as linear combinations of the $\mathbb{Z}$ translates of $\phi_H$ and $\phi_S$. This leads to the following definition.

$(\boldsymbol{iv})$ **Definition.** A square integrable function $\phi$: $\mathbb{R}^n \mapsto \mathbb{C}$ is a $\mathbb{Z}^n$- sampling function if the following hold:

(S1) $\phi$ is bounded on $\mathbb{R}^n$ and is continuous on a dense open subset $\mathcal{U} \subset \mathbb{R}^n$ ($e.g.$, $\mathcal{U}$ might be the complement of finitely many smooth surfaces having dimensions $<n$) ;

(S2) $\mathcal{B}_\phi = \{\mathrm{T}_k\phi : k \in \mathbb{Z}^n\}$ is a frame for $<\phi>_{\mathbb{Z}^n}$, $i.e.$, there are positive constants A and B for which

$$\mathrm{A}\|f\|^2_{L^2(\mathbb{R}^n)} \leq \sum_{k\in\mathbb{Z}^n} |< f, \mathrm{T}_k\phi >|^2 \leq \mathrm{B}\|f\|^2_{L^2(\mathbb{R}^n)} \quad (16)$$

for all $f \in <\phi>_{\mathbb{Z}^n}$

(S3) Changing notation, for each of the $a.e.$ equivalence classes comprising $<\phi>_{\mathbb{Z}^n}$ we can pick a specific representative $f$ in the class such that, for each $x \in \mathbb{R}^n$,

$\sum_{k\in\mathbb{Z}^n} f(k)\phi(x-k)$ converges unconditionally to $f(x)$ and the convergence is uniform on compact subsets of $\mathcal{U}$ so $f$ is continuous on $\mathcal{U}$.

"Essentially", one can show that, when (S1) and (S2) hold, then (S3) holds $\Leftrightarrow$ $Z\phi(0,\ \cdot\ ) = \chi_{supp\,[\phi,\phi]_{\mathbb{Z}^n}}\ a.e.$ We will not bother to discuss the added technical conditions needed to make this statement correct nor will we take time to go into the details of the proof. In an obvious way, one can define Zak transforms with an arbitrary lattice $\mathcal{L} \subset \mathbb{R}^n$ replacing $\mathbb{Z}^n$ and use them to characterize square integrable $\mathcal{L}$ -sampling functions.

There are also sampling functions where, as with $\phi_H$, the sample set need not be the points on a lattice; these are closely linked to the theory of reproducing kernel spaces. Finally, one can forget about square integrability and look just at functions satisfying a version of (S3). It should be clear that the subject of sampling functions is of great interest in applications since members $f$ of a sampling space determined by a sampling function $\phi$ are uniquely determined by their values on a countable sample set and there is an explicit way to reconstruct each $f$ from its sample values and certain translates of $\phi$.

**3.7. Summary.** As mentioned in 3.1, all of the standard reproducing function systems, including Gabor and wavelet systems as well as other more general systems, rely on lattice translations of certain generating functions. In this sense, the theory of shift invariant spaces is fundamental for construction and implementation of reproducing function systems. We have discussed all of the basic ingredients of this theory, culminating in the very important Dimension Theorem and a preliminary discussion of $L^2$ sampling functions We could go on at great length to develop more ingredients of shift invariant space theory and sampling theory but it's better to develop new ingredients as they are needed for specific problems and applications. In effect, every researcher in applied harmonic analysis needs to know the basic ingredients in order to apply them to his/her research agenda and, as necessary, prove new theorems and develop new computational techniques. It's not an exaggeration to say that, in the same way that multi-variable calculus is impossible to fully understand without linear algebra, modern applied harmonic analysis is impossible to fully understand without shift invariant space theory.

## 4. Applications of Shift Invariant Space Theory to Wavelet Systems

**4.1.** Let us return to the discussion in Section **2** of dyadic ON MRAs $(V_j)_{j\in\mathbb{Z}}$ for $L^2(\mathbb{R})$, focusing our attention on the $\mathbb{Z}$ shift – invariant space $V_0 = \ <\phi>_{\mathbb{Z}}$, the dilated space $V_1 = D(V_0)$ with $D = D_{1/2}$ , and the orthogonal complement $W_0$ of $V_0$ in $V_1$. Recall that $V_0 \subset V_1$ by MRA(1) and, by MRA(4), $\mathcal{B}_{\phi,\mathbb{Z}} = \{T_k\phi : k \in \mathbb{Z}\}$ is an orthonormal basis of $V_0$. From Section **3**, this tells us that $[\phi,\phi]_{\mathbb{Z}} = 1 = \boldsymbol{dim}_{V_0,\mathbb{Z}}\ a.e.$ Since $DT_k\phi = T_{k/2}D\phi$ and D is a unitary map, $D\,\mathcal{B}_{\phi,\mathbb{Z}}=\mathcal{B}_{D\phi,\frac{1}{2}\mathbb{Z}}$ is an orthonormal basis for $V_1 = \ < D\phi >_{\frac{1}{2}\mathbb{Z}}$. On the other hand, $\mathbb{Z} \subset \frac{1}{2}\mathbb{Z}$ so $V_1$ is also a $\mathbb{Z}$ shift-invariant space. Because $DT_{2j}\phi = T_jD\phi$ and $DT_{2j+1}\phi = T_jDT_1\phi$, $\mathcal{B}_{D\phi,\frac{1}{2}\mathbb{Z}}$ is the union of the orthonormal basis $\mathcal{B}_{D\phi,\mathbb{Z}}$ for $< D\phi >_{\mathbb{Z}}$ and the orthonormal basis $\mathcal{B}_{DT_1\phi,\mathbb{Z}}$ for $< DT_1\phi >_{\mathbb{Z}}$. Also $V_1$ is the orthogonal direct sum of these two $\mathbb{Z}$ shift invariant spaces so $\boldsymbol{dim}_{V_1,\mathbb{Z}}=2\ a.e.$ and we read off from Corollary 3.5 that $\boldsymbol{dim}_{W_0,\mathbb{Z}} = \boldsymbol{dim}_{V_1,\mathbb{Z}} - \boldsymbol{dim}_{V_0,\mathbb{Z}} = \mathbf{1}\ a.e.$ As we claimed in Section **2.9,** this means there is a function $\psi \in W_0$ for which $\mathcal{B}_{\psi,\mathbb{Z}}$ is an orthonormal basis for $W_0$. From MRA(1) – MRA(4), it follows that $L^2(\mathbb{R}) = \bigoplus_{j\in\mathbb{Z}} D^jW_0$ and, because D is unitary

$\{\psi_{j,k} = \mathrm{D}^j\mathrm{T}_k\psi : (j,k) \in \mathbb{Z}\times\mathbb{Z}\}$ is an orthonormal basis for $\mathrm{L}^2(\mathbb{R})$. Note that $\psi$ is unique up to a unimodular multiplier in the sense that, for each $f \in \mathrm{W}_0$, $\widehat{f} = m\widehat{\psi}$ with $m = [f,\psi]_{\mathbb{Z}}$ in $\mathrm{L}^2(\widehat{\mathbb{R}}/\mathbb{Z})$ and $[f,f]_{\mathbb{Z}} = |m|^2[\psi,\psi]_{\mathbb{Z}} = |m|^2$ so $\mathcal{B}_{f,\mathbb{Z}}$ is an orthonormal basis for $\mathrm{W}_0 \Leftrightarrow 1 = [f,f]_{\mathbb{Z}} = |m|^2\ a.e.$

**4.2**. Suppose we start with $\phi \in \mathrm{L}^2(\mathbb{R}^n), n \geq 2$ for, which $[\phi,\phi]_{\mathbb{Z}^n} = 1\ a.e.$ and define $\mathrm{V}_0$ to be $<\phi>_{\mathbb{Z}^n}$. Also suppose we have $a \in \mathrm{GL}(n,\mathbb{Z}) = \{$members of $\mathrm{GL}(n,\mathbb{R})$ with integer entries$\}$ for which $\mathrm{V}_0 \subset \mathrm{D}_{a^{-1}}\mathrm{V}_0 \equiv \mathrm{V}_1$. As in 4.1, $\mathrm{V}_1$ is a principal $a^{-1}\mathbb{Z}^n$ shift invariant space with the orthonormal basis $\mathcal{B}_{\mathrm{D}_{a^{-1}}\phi, a^{-1}\mathbb{Z}^n}$. By a standard result in linear algebra, $\mathbb{Z}^n \subset a^{-1}\mathbb{Z}^n$ and the quotient group $a^{-1}\mathbb{Z}^n/\mathbb{Z}^n$ is finite with $|\det a|$ members. Just as in 4.1, we deduce that, for $\mathrm{W}_0$ the orthogonal complement of $\mathrm{V}_0$ in $\mathrm{V}_1$,

$$\boldsymbol{dim}_{\mathrm{W}_0,\mathbb{Z}^n} = \boldsymbol{dim}_{\mathrm{V}_1,\mathbb{Z}^n} - \boldsymbol{dim}_{\mathrm{V}_0,\mathbb{Z}^n} = |det\, a| - 1\ a.e.$$

so, with N=$|\det a|$ there exist choices of OPFGS sets $\Psi=\{\psi^{(1)},\cdots\psi^{(\mathrm{N}-1)}\}$ for $\mathrm{W}_0$ whose $\mathbb{Z}^n$ translates form an orthonormal basis for $\mathrm{W}_0$. There is a theorem giving necessary and sufficient conditions on $\phi$ and $a$ in order that we have the $\mathbb{R}^n$analog of the MRA(2) union property and it turns out that the analog of the MRA(3) intersection property automatically holds. When these conditions on $\phi$ and $a$ are satisfied, defining $\mathrm{V}_j$ to be $\mathrm{D}_{a^{-j}}\mathrm{V}_0$ for $j \in \mathbb{Z}$, $\mathrm{W}_j \equiv \mathrm{D}_{a^{-j}}\mathrm{W}_0$ is the orthogonal complement of $\mathrm{V}_j$ in $\mathrm{V}_{j+1}$, $\mathrm{L}^2(\mathbb{R}^n) = \bigoplus_{j\in\mathbb{Z}}\mathrm{W}_j$, and the multi-generated wavelet system $\{\mathrm{D}_{a^{-j}}\mathrm{T}_k\psi^{(l)} : j \in \mathbb{Z},\ k \in \mathbb{Z}^n, 1 \leq l \leq \mathrm{N}\text{-}1\}$ is an orthonormal basis for $\mathrm{L}^2(\mathbb{R}^n)$. For obvious reasons, we say all of the wavelet systems obtained from some choice of $\Psi$ are <u>associated</u> with the MRA $(\mathrm{V}_j)_{j\in\mathbb{Z}}$ and the <u>scaling function</u> $\phi$ for this MRA. All of this is fine when N=2 and we then have only a singly generated system with $\psi$=$\psi^{(1)}$ uniquely determined by $\phi$ modulo a unimodular multiplier. It's not so good when N>2.

For one thing, there's no canonical choice for $\Psi$ and any two choices $\Psi$, $\Psi'$ are related by an (N-1) $\times$ (N-1) matrix of brackets which, for *a.e.* $\xi$, is unitary. Even if one gets around this difficulty by using special properties of $\phi$ to obtain what seems to be a nice choice for $\Psi$, the resulting reproducing function system is very likely to be inefficient, *i.e.*, many coefficients need to be saved in order to obtain reasonable good approximations for signals.

**4.3 Example: The two-dimensional dyadic Haar system.**
Recall that the 1-dimensional Haar scaling function is $\phi_H = \chi_{[0,1)}$ and the associated 1-dimensional Haar wavelet function is $\psi_{\mathrm{H}} = \chi_{[0,\frac{1}{2})} - \chi_{[\frac{1}{2},1)}$. When $f$ and $g$ are functions on $\mathbb{R}$, $f \otimes g$ denotes the function on $\mathbb{R}^2$ whose value at $(x, y)$ is $f(x)g(y)$. In particular, $\phi_H^{(2)} = \phi_H \otimes \phi_H = \chi_{[0,1)^2}$ and $\mathrm{V}_0 = \ < \phi_H^{(2)} >_{\mathbb{Z}^2}$ has the $\mathbb{Z}^2$ translates of $\phi_H^{(2)}$ as an orthonormal basis and consists of the functions in $\mathrm{L}^2(\mathbb{R}^2)$ which are constant on each $\mathbb{Z}^2$-translate of the square $[0,1)^2$. The dyadic dilation operator in two dimensions is $\mathrm{D}=\mathrm{D}_{\begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}}$ with $\mathrm{V}_0 \subset \mathrm{V}_1 \equiv \mathrm{D}(\mathrm{V}_0)$. From our discussion above, $\boldsymbol{dim}_{W_0,\mathbb{Z}^2} =$ $=4$-$1$=$3$ *a.e.* One can easily check that a choice for $\Psi$ is $\{\phi_H \otimes \psi_H, \psi_H \otimes \phi_H, \psi_H \otimes \psi_H\}$. Along all of the boundary edges for the 4 squares of side 1/2 which comprise $[0,1)^2$, we have discontinuities for the scaling function $\phi_H^{(2)}$ and the members of $\Psi$. Certainly these discontinuities lower the efficiency of the system. But, even if we replace $\phi_H$ and $\psi_H$ by one of the Daubechies scaling functions and associated wavelet function, it turns we still have resolution problems along the line $y = x$ and these resolution problems are just unavoidable with tensor product wavelets.

**4.4. The twin-dragon example.** For wavelet purposes, one of the easiest $2 \times 2$ integer matrices with determinant 2 is the quincunx matrix $q = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$. Taking D to be $\mathrm{D}_{q^{-1}}$, it's easy to construct Shannon-like wavelet systems using dilations by powers of D and the lattice $\mathbb{Z}^2$. But, as Gröchenig and Madych showed, the only Haar-like wavelet system using these operators has $\psi$ the difference of the characteristic functions of two congruent parts of a fractal set called the twin dragon. [Likely, a Googol search with the code phrase twin dragon will identify some websites showing pictures of the twin dragon and, perhaps, some other wavelets supported on fractal sets.] It's unknown whether there any Daubechies-like systems using just dilations by powers of $q$ and lattice translations. This is more evidence that wavelet theory in higher dimensions is much more difficult than the one-dimensional theory.

**4.5. Composite wavelets.** The difficulties discussed above with higher dimensional wavelet systems using only lattice translations and integer powers of a fixed matrix $a$ provided part of the motivation for our research group at Washington University in St. Louis to introduce a new type of wavelet systems which we call composite wavelets. These have the form $\{\mathrm{D}_a^i \mathrm{D}_b \mathrm{T}_l \psi : i \in \mathbb{Z}, b \in \mathrm{B}, l \in \mathcal{L}\}$ where $a \in \mathrm{GL}(n,\mathbb{R})$, $\mathcal{L}$ is a lattice in $\mathbb{R}^n$, and B is a group of matrices taking $\mathcal{L}$ onto $\mathcal{L}$. In particular, $|\det b| = 1$ for all $b$.

$(i)$ When $\mathcal{L} = \mathbb{Z}^2$, $a = \begin{pmatrix} c & 0 \\ 0 & \sqrt{c} \end{pmatrix}$ for some $c > 1$,

and $\mathrm{B} = \left\{ \begin{pmatrix} 1 & j \\ 0 & 1 \end{pmatrix} : j \in \mathbb{Z} \right\}$, the resulting wavelet systems are called shearlets since the action on $\mathbb{R}^2$ of the members of B is by shearing transformations (the terminology comes from an analog of wind shearing forces in physics). One can construct orthonormal MRA shearlets with multiple generators but much better is to constract Parseval frame MRA shearlets with a single generator. Along with somewhat similar but more complicated systems called curvelets (invented by Candes and Donoho), shearlets have been recently shown to have close to optimal efficiency for two-dimensional signals ("photographs" of two-dimensional objects) having discontinuities only along a finite number of $C^2$ curves (intuitively, the edges of the objects in the photograph). In particular, curvelets and shearlets out perform all other known reproducing function systems for signals of this type.

$(ii)$ When $\mathcal{L}$=$\mathbb{Z}^2$, $a = q$, and B is the 8-element group of symmetries of the unit square (4 rotations by integer multiples of $\pi/2$ plus 4 orthogonal reflections), our group showed that there is a very easy orthonormal Haar-like MRA composite wavelet system with $\phi$ a constant times the characteristic function of an isosceles right triangle and $\psi$ the difference of the characteristic functions of two congruent subtriangles. Obviously, this is a huge improvement over the twin dragon example. Our former Ph.D. student J. Blanchard and his collaborator K. Steffen recently generalized our result to 11 of the famous 17 crytallographic groups of plane rotations and reflections. There is evidence that it may be possible to replace characteristic functions with compactly supported functions having a certain amount of smoothness, hence analogous to the Daubechies wavelets in one dimension.

$(iii)$ There are many open questions concerning composite wavelets. Because of the success mentioned in $(i)$ and $(ii)$, composite wavelets appear to be very promising as an address to the problems mentioned previously for higher dimensional wavelet theory.

## 4.6. Low and high pass filters, Smith-Barnwell equations, and the filter bank technique for dyadic ON MRAs

Let's return to the 2.10 and 4.1 context of a dyadic ON MRA $(\mathrm{V}_j)_{j\in\mathbb{Z}} = (\mathrm{D}^j\mathrm{V}_0)_{j\in\mathbb{Z}}$ for $\mathrm{L}^2(\mathbb{R})$ with scaling function $\phi$ and associated wavelet function $\psi$. Recall that $(\mathrm{D}f)(x) = \sqrt{2}\,f(2x)$, $\mathrm{V}_1=\mathrm{V}_0 \oplus \mathrm{W}_0$ with $\{\mathrm{DT}_k\phi : k \in \mathbb{Z}\}$ an orthonormal basis for $\mathrm{V}_1$, $\{\mathrm{T}_k\phi : k \in \mathbb{Z}\}$ an orthonormal basis for $\mathrm{V}_0$, and $\{\mathrm{T}_k\psi : k \in \mathbb{Z}\}$ an orthornormal basis for $\mathrm{W}_0$.

$(i)$ It's convenient to use instead the orthogonal basis for $\mathrm{V}_1$ consisting of the functions $\sqrt{2}\,(\mathrm{DT}_{-k}\phi)(x) = 2\phi(2x+2k)$ and to similarly replace $k$ with $-k$ in the bases for $\mathrm{V}_0$ and $\mathrm{W}_0$. For each $f \in \mathrm{V}_1$, we then have square summable coefficient sequences $(a_k)_{k\in\mathbb{Z}}$, $(b_k)_{k\in\mathbb{Z}}$, and $(c_k)_{k\in\mathbb{Z}}$ uniquely determined by $f$ such that, in the sense of $\mathrm{L}^2(\mathbb{R})$ convergence,

$$f(x) = \sum_{k\in\mathbb{Z}} a_k\phi(x+k) + \sum_{k\in\mathbb{Z}} b_k\psi(x+k)$$

$$= 2\sum_{k\in\mathbb{Z}} c_k\phi(2x+k) \qquad (12)$$

It's convenient to replace $x$ by $x/2$ in the second equation in (12) and to divide both sides by $\frac{1}{2}$, thereby recasting this equation as

$$\tfrac{1}{2}f(x/2) = \sum_{k\in\mathbb{Z}} c_k\phi(x+k) \qquad (13)$$

[For the special cases $f = \phi$ and $f = \psi$, the two equations of the form (13) are called the refinement equations for $\phi$ and $\psi$.]

Now let $p_0(\xi) = \sum_{k\in\mathbb{Z}} a_k e_k(\xi)$, $q(\xi) = \sum_{k\in\mathbb{Z}} b_k e_k(\xi)$, and $p_1(\xi) = \sum_{k\in\mathbb{Z}} c_k e_k(\xi)$ be the members of $L^2(\widehat{\mathbb{R}}/\mathbb{Z})$ whose Fourier series expressions are given by our three coefficient sequences. By applying the Fourier transform $\mathcal{F}$ to both sides of the top equation in (12) and to both sides of (13), a routine change of variable computation gives

$$\widehat{f}(2\xi) = p_1(\xi)\widehat{\phi}(\xi) = p_0(2\xi)\widehat{\phi}(2\xi) + q(2\xi)\,\widehat{\psi}(2\xi) \;\; a.e. \quad (14)$$

When each of $\phi$,$\psi$, and $f$ have compact support (assumed in many applications), we obviously have only finitely many non-zero coefficients in each of the sums in (12) and the functions $p_1$, $p_0$, and $q$ are trig polynomials. As we'll see below, (14) is the key to the very fast filter bank technique for calculating the $a_k's$ and $b_k's$ in terms of the $c_k's$ and this technique doesn't need closed-form expressions (formulas) for $\phi$ and $\psi$, only that $\phi$ and $\psi$ are known to exist! In fact, for the very popular Daubechies systems, we don't have closed-form expressions for $\phi$ and $\psi$ and it's very dubious whether such expressions will ever be found.

$(ii)$ **Definition.** The low pass filter $m_0$ and high pass filter $m_1$ determined by $\phi$ and $\psi$ are the members of $L^2(\widehat{\mathbb{R}}/\mathbb{Z})$ for which

$$\widehat{\phi}(2\xi) = m_0(\xi)\widehat{\phi}(\xi)\, a.e., \quad \widehat{\psi}(2\xi) = m_1(\xi)\widehat{\phi}(\xi)\, a.e. \qquad (15)$$

The existence of these filters is provided by the first part of (14) for the special cases $f = \phi$ and $f = \psi$.

$(iii)$ **Lemma : Using the above notations, $m_0$ and $m_1$**

**satisfy the three dyadic Smith-Barnwell equations necessary and sufficient to have M($\xi$)=$\begin{pmatrix} m_0(\xi) & m_0(\xi+\frac{1}{2}) \\ m_1(\xi) & m_1(\xi+1/2) \end{pmatrix}$ be a unitary matrix for $a.e.\ \xi$.**

**Proof.** The dyadic Smith-Barnwell equations arise by remembering that a $2 \times 2$ complex matrix is unitary precisely when its rows are perpendicular unit vectors. Thus, the three equations are:

$$\begin{aligned} 0 &= |m_0(\xi)|^2 + |m_0(\xi{+}1/2)|^2 - 1 \\ &= |m_1(\xi)|^2 + |m_1(\xi{+}1/2)|^2 - 1 \\ &= (m_0\overline{m_1})(\xi) + (m_0\overline{m_1})(\xi + 1/2)\,. \end{aligned} \qquad (16)$$

Recall that, for $f$ and $g$ in $L^2(\mathbb{R})$, $[f, g]_{\mathbb{Z}}(\xi)$ is the sum over $k \in \mathbb{Z}$ of the terms $(\widehat{f}\,\overline{\widehat{g}})(\xi + k)$. We can then break this up into the sum over the even integers $k = 2j$ and the sum over the odd integers $k = 2j + 1$. For convenience, we can also replace the variable $\xi$ with $2\xi$. Taking separately the 3 cases $(f, g) = (\phi, \phi)$, $(\psi, \psi)$, or $(\phi,\psi)$ and using (15) along with $[\phi, \phi]_{\mathbb{Z}}=[\psi,\psi]_{\mathbb{Z}} = 1\ a.e.$, $[\phi, \psi]_{\mathbb{Z}} = 0\ a.e.$ (by the 3.3 properties for brackets), an easy computation shows (16) holds for $a.e.\ \xi$.

$(iv)$ **Filter Bank Technique.** Let us return to (14) in 4.6$(i)$. Using (15), the uniqueness of periodic multipliers for members of shift invariant spaces tells us that

$$p_1(\xi) = p_0(2\xi)m_0(\xi) + q(2\xi)\, m_1(\xi)\ a.e. \qquad (17)$$

When we replace the variable $\xi$ in (17) with $\xi$+1/2, $2\xi$ is replaced by $2\xi$+1. However, since $p_0$ and $q$ are $\mathbb{Z}$-periodic , $p_0(2\xi+1)=p_0(2\xi)$ and similarly for $q$. This allows us to turn (17) into the matrix equation

$$\begin{pmatrix} p_1(\xi) & p_1(\xi+1/2) \end{pmatrix} = \begin{pmatrix} p_0(2\xi) & q(2\xi) \end{pmatrix} M(\xi) \text{ a.e.} \quad (18)$$

for $p_0$ and $q$ which we solve simply by multiplying both sides by $M(\xi)^{-1} = M(\xi)^* = \begin{pmatrix} \overline{m_0(\xi)} & \overline{m_1(\xi)} \\ \overline{m_0(\xi+1/2)} & \overline{m_1(\xi+1/2)} \end{pmatrix}$.
This determines the Fourier coefficients $(a_k)_{k\in\mathbb{Z}}$ for $p_0$ and $(b_k)_{k\in\mathbb{Z}}$ for $q$ in terms of the Fourier coefficients $(c_k)_{k\in\mathbb{Z}}$ for $p_1$ and the Fourier coefficients for $m_0$ and $m_1$. When each of $p_1$,$m_0$, and $m_1$ is a trig polynomial (has only finitely many non-zero coefficients), $p_0$ and $q$ also are trig polynomials. In the language used in 2.10, we have just explained how the filter bank technique converts the coefficients expressing a first resolution level approximation $f_1$ in terms of the basis for $V_1$ into the coefficients needed to express the zeroth level approximation $f_0$ and the error term $E_0 = f_1 - f_0$. For any $i \geq 2$, the $i^{th}$ resolution level approximation $f_i \in V_i$ for a signal $f \in V_j$, $j \geq i$, is just the image under $D^{i-1}$ of a member $g_1$ of $V_1$. Using our technique to convert coefficients $g_1$ to coefficients for $g_0 \in V_0$ and for $g_1 - g_0 \in W_0$ is equivalent to converting coefficients for $f_i$ to coefficients for $f_{i-1} \in V_{i-1}$ and $E_{i-1} = f_i - f_{i-1}$.
There are variations on the filterbank technique in higher dimensions with larger matrices and for situations where $\phi$ and $\psi$ aren't compactly supported but vanish very rapidly at $\pm\infty$ where truncation methods are used which, in practice, don't significantly impair resolution. Also, with some technical fussing, the filter bank technique adapts to the case of dyadic Parseval frame MRA scaling functions $\phi$ and associated dyadic Parseval frame MRA wavelets $\psi$, *i.e.* $\{T_k\phi : k \in \mathbb{Z}\}$ is a Parseval frame for $V_0$ and $\{T_k\psi : k \in \mathbb{Z}\}$ is a Parseval frame for $W_0$. Very likely, the magic of filter banks was as big a factor as increased efficiency in the decision arrived at in the late 1990's to make wavelets the industrial standard for data analysis.

### 4.7 Construction of dyadic MRAs

Reversing the arguments in 4.6$(iii)$, suppose we start with a $\mathbb{Z}$-periodic function $m_0$ satisfying the first of the 3 equations in (16). Defining $m_1$ by $m_1(\xi) = e^{2\pi i\xi} m_0(\xi + 1/2)$, it's simple to check that all three of the equations in (16) are satisfied. There's no guarantee that there exists a function $\phi$ satisfying the first of the equations in (15), but, if it exists, we can define $\psi$ by insisting that the second equation in (15) holds. It's not hard to see that the only "reasonable" candidate for $\phi$ is given by $\widehat{\phi}(\xi) = \prod_{j=1}^{\infty} m_0(\xi/2^j)$ so we have to limit attention to choices for $m_0$ not only giving $a.e.$ convergence of this infinite product but also yielding $[\phi,\phi]_{\mathbb{Z}} = \chi_\Omega \; a.e.$ for some subset $\Omega$ of $\widehat{\mathbb{R}}/\mathbb{Z}$. The point of this very round-about procedure is that all dyadic Parseval frame MRA wavelets arise this way with $\widehat{\phi}$ satisfying a technical condition called dyadic continuity at 0 which turns out to be necessary and sufficient for MRA(2). Moreover, it was by exactly this procedure that Daubechies used to demonstrate the existence of her compactly supported orthonormal scaling and wavelet functions when $m_0$ is taken to be one of a very special class of trig polynomials.

## 5. Continuous Wavelet Transforms and the Unified Theorem

### 5.1 Continuous Translation Systems.

***(i)*** For $f$ and $\psi$ in $L^2(\mathbb{R}^n)$ with $\psi^*(x) \equiv \overline{\psi(-x)}$ and $g(y) = \langle f, T_y\psi \rangle$, the definition of the convolution operation gives $g = f*(\psi^*)$ and it follows that $\widehat{g} = \widehat{f}\,\overline{\widehat{\psi}}$ so

$$||g||^2_{L^2(\mathbb{R}^n)} = ||\widehat{g}||^2_{L^2((\mathbb{R}^n)^\frown)} = \int_{(\mathbb{R}^n)^\frown} |\widehat{f}|^2 |\widehat{\psi}|^2 d\xi\,. \qquad (1)$$

Note that $g$ is a bounded, continuous function for every choice of $f$ and $\psi$. For $n = 1$ and N $> 1$, $\sum\limits_{k\in\mathbb{Z}} \frac{1}{N} |< f, \mathrm{T}_{k/N}\psi >|^2$ is a Riemann sum approximation with increments of 1/N for $||g||^2_{L^2(\mathbb{R}^n)}$. If $g$ happens to be uniformly continuous, the limit of these Riemann sum approximations as N→∞ is equal to $||g||^2_{L^2(\mathbb{R}^n)}$. On the other hand, the N$^{th}$ approximation is a shift invariant space expression for the lattice $\frac{1}{N}\mathbb{Z}$ and, from 3.2 is equal to 1/ N$||[f, \psi]||^2_{L^2(\widehat{\mathbb{R}}/N\mathbb{Z})}$ =1/N$\int_0^N |[f, \psi]|^2$. We can eleminate the 1/N factors by replacing $\psi$ with $\phi^{(\mathrm{N})} = \psi/\sqrt{\mathrm{N}}$ and then $\psi = \sqrt{\mathrm{N}}\,\phi^{(\mathrm{N})}$ . Obviously 1/N is the length of every $\frac{1}{N}\mathbb{Z}$ tiling domain and N is the reciprocal of 1/N. For $\mathcal{L} = a\mathbb{Z}^n$ a lattice in $\mathbb{R}^n$ we denote by $|\mathcal{L}| \equiv |\det a|$ the Lebesgue measure of every $\mathcal{L}$ – tiling domain. The above remarks suggest the properties of an $\mathcal{L}$ shift-invariant space system $\{\mathrm{T}_l\phi : l \in \mathcal{L}\}$ should be compared with those of the continuous translation system $\{\mathrm{T}_y\,(\frac{\phi}{|\mathcal{L}|^{1/2}}) : y \in \mathbb{R}^n\}$ . Below, we'll state a very general theorem that illustrates this rescaling principle. But first, let's look at some examples of continuous translation systems.

$(ii)$ **Definition.** Suppose $I$ is a countable index set and $\Psi = \{\psi_i : i \in I\}$ is a subset of $\mathrm{L}^2(\mathbb{R}^n)$. Then the $\mathbb{R}^n$ -translates of the members of $\Psi$ are a continuous reproducing function system for $\mathrm{L}^2(\mathbb{R}^n)$ if, for every $f \in \mathrm{L}^2(\mathbb{R}^n)$,

$$||f||^2_{L^2(\mathbb{R}^n)} = \sum_{i\in I} \int_{\mathbb{R}^n} |< f, \mathrm{T}_y\psi_i >|^2 dy \qquad (2)$$

Using the polarization identity, (2) is equivalent to

$$< f,g>_{L^2(\mathbb{R}^n)}=\sum_{i\in I}\int_{\mathbb{R}^n} < f, \mathrm{T}_y\psi_i > <\mathrm{T}_y\psi_i, g > dy \qquad (3)$$

$\forall f, g \in \mathrm{L}^2(\mathbb{R}^n)$ and it's customary to describe (3) by saying that the continuous reproducing formula

$$f(x) = \sum_{i\in I}\int_{\mathbb{R}^n} < f, \mathrm{T}_y\psi_i > \psi_i(x-y)dy \qquad (4)$$

holds in the weak sense.

In view of (1), elementary measure theory shows that (2) holds $\Leftrightarrow$

$$\sum_{i\in I}|\widehat{\psi}_i(\xi)|^2 = 1 \text{ for } a.e. \xi \in (\mathbb{R}^n)\widehat{\ } \qquad (5)$$

For reasons we'll explain below, (5) is said to be an example of a Calderón equation.

$(iii)$ **Definition.** Suppose $A = \{a_i : i \in I\}$ is a countable subset of $\mathrm{GL}(n, \mathbb{R})$ and $\psi \in \mathrm{L}^2(\mathbb{R}^n)$. Then $\{\mathrm{D}_{a_i}\mathrm{T}_y\psi : y \in \mathbb{R}^n\}$ is a continuous Parseval frame wavelet system for $\mathrm{L}^2(\mathbb{R}^n)$ with discrete $A$-dilations if

$$||f||^2_{L^2(\mathbb{R}^n)}=\sum_{i\in I}\int_{\mathbb{R}^n}| < f, \mathrm{D}_{a_i}\mathrm{T}_y\psi > |^2 dy \quad \forall f \in \mathrm{L}^2(\mathbb{R}^n) \quad (6)$$

Using $\mathrm{D}_{a_i}\mathrm{T}_y = \mathrm{T}_{a_i y}\mathrm{D}_{a_i}$, we can, for each $i$, introduce the change of integration variable $y' = a_i y$ or $y = a_i^{-1} y'$. Then, by (1) and the formula $(\mathrm{D}_{a_i}\psi)\widehat{\ }(\xi) = |\det a_i|^{1/2}\widehat{\psi}(\xi a_i)$ from Theorem 1.3,

$$\int_{\mathbb{R}^n}| < f, \mathrm{D}_{a_i}\mathrm{T}_y\psi > |^2 dy = \int_{\mathbb{R}^n}| < f, \mathrm{T}_{y'}\mathrm{D}_{a_i}\psi > |^2 \frac{1}{|det\, a_i|} dy'$$

$$= \int_{(\mathbb{R}^n)^\frown} |\widehat{f}(\xi)|^2 |\widehat{\psi}(\xi a_i)|^2 d\xi \qquad (7)$$

Using (7) for each $i \in I$, we deduce as in $(ii)$ that (6) holds $\Leftrightarrow$ $\psi$ satisfies the discrete Calderón equation.

$$\sum_{i \in I} |\widehat{\psi}(\xi a_i)|^2 = 1 \ a.e. \qquad (8)$$

## 5.2 Admissible Affine Groups

**$(i)$ Definitions:** The full affine group on $\mathbb{R}^n$ is the group of transformations on $\mathbb{R}^n$ of the form $x \mapsto (a, y) \cdot x \equiv a(x + y)$ for $(a, y) \in \mathrm{GL}(n, \mathbb{R}) \times \mathbb{R}^n$. It's convenient to denote this group by $\mathrm{Aff}(n, \mathbb{R}) = \{g = (a, y) : (a, y) \in \mathrm{GL}(n, \mathbb{R}) \times \mathbb{R}^n\}$ with the group multiplication law on $\mathrm{Aff}(n, \mathbb{R})$ determined by insisting that $g_1 \cdot (g_2 \cdot x) = (g_1 g_2) \cdot x$ for all $(g_1, g_2, x) \in \mathrm{Aff}(n, \mathbb{R}) \times \mathrm{Aff}(n, \mathbb{R}) \times \mathbb{R}^n$. Then, with $id_n$ the $n \times n$ identity matrix, $a \mapsto (a, 0)$ is an isomorphism from $\mathrm{GL}(n, \mathbb{R})$ onto a subgroup of $\mathrm{Aff}(n, \mathbb{R})$ and $y \mapsto (id_n, y)$ is an isomorphism from $\mathbb{R}^n$onto a normal subgroup of $\mathrm{Aff}(n, \mathbb{R})$. In this sense, $\mathrm{Aff}(n, \mathbb{R})$ is the semi-direct product of $\mathrm{GL}(n, \mathbb{R})$ and $\mathbb{R}^n$. By an affine group, we mean a subgroup G of $\mathrm{Aff}(n, \mathbb{R})$ of the form $\{(a, y) : a \in A,\ y \in \mathbb{R}^n\}$ where $A$ is a not necessarily connected Lie subgroup of $\mathrm{GL}(n, \mathbb{R})$, $e.g.$, any closed subgroup of $\mathrm{GL}(n, \mathbb{R})$. Then G is the semi-direct product of $A$ and $\mathbb{R}^n$. In this case, for any left Haar measure $\mu$ on $A$, we have an associated left Haar measue $\nu$ on G defined by $d\nu(a, y) = d\mu(a)dy$ [Recall that left Haar measures on topological groups are Borel measures invariant under left translation and any two left Haar measures on a group are positive scalar multiples of each other].

***(ii)* More definitions.** Suppose G is as in $(i)$. Then $(a, y) \mapsto \tau_{(a,y)} \equiv \mathrm{D}_a \mathrm{T}_y$ defines a unitary representation $\tau$ of G on $\mathrm{L}^2(\mathbb{R}^n)$ [Group theorists call $\tau$ the quasi left regular representation of G acting on square integrable functions on the homogeneous space $\mathbb{R}^n \approx A\backslash G$].

An admissible vector for $\tau$ is a non-zero member $\psi$ of $\mathrm{L}^2(\mathbb{R}^n)$ such that, with

$$(\mathrm{W}_\psi f)(a, y) = \, < f, \tau_{(a,y)}\psi > ,$$

$\mathrm{W}_\psi$ is an isometry from $\mathrm{L}^2(\mathbb{R}^n)$ onto a closed subspace of $\mathrm{L}^2(G, \nu)$. [Group theorists say that $\tau$ is square integrable if admissible vectors exist]. Unraveling this, $\psi$ is admissible for $\tau \Leftrightarrow \forall f \in \mathrm{L}^2(\mathbb{R}^n)$, we have

$$||f||^2_{L^2(\mathbb{R}^n)} = \int_A \left(\int_{\mathbb{R}^n} | < f, \mathrm{D}_a \mathrm{T}_y \psi > |^2 \, dy\right) d\mu(a) \tag{9}$$

***(iii)* Theorem.** Using the definitions and notations in $(i)$ and $(ii)$, $\psi$ is an admissible vector for $\tau \Leftrightarrow$ the orbit integral

$$\sigma_\psi(\xi) \equiv \int_A |\widehat{\psi}(\xi a)|^2 d\mu(a)$$

satisfies the Calderón equation $\sigma_\psi = 1$ $a.e.$

**Proof.** We merely need to apply, for each $a \in A$, the same trick that we used above in 5.1$(iii)$, namely, $\mathrm{D}_a \mathrm{T}_y = \mathrm{T}_{ay} \mathrm{D}_a$ so the change of variable $y' = ay$ in (9) makes the determinant factor in $(\mathrm{D}_a \psi)\hat{}$ disappear and the result is that (9) holds for each $f \Leftrightarrow$

$$||f||^2_{L^2(\mathbb{R}^n)} = \int_{(\mathbb{R}^n)\hat{}} |\widehat{f}(\xi)|^2 d\xi = \int_{(\mathbb{R}^n)\hat{}} |\widehat{f}(\xi)|^2 \sigma_\psi(\xi) d\xi \tag{10}$$

By elementary measure theory, the second equation in (10) holds $\forall \widehat{f} \in \mathrm{L}^2((\mathbb{R}^n)\hat{}) \Leftrightarrow \sigma_\psi = 1$ $a.e.$

**Note:** Because $\mu$ is left invariant, $\sigma_\psi(\xi b) = \sigma_\psi(\xi)$ for each $\xi$ and each $b \in A$. Hence $\sigma_\psi$ is constant on $A$ – orbits. In the early 1960's, Alberto Calderón devised the above proof for the case

$n = 1$ with $A =$ GL(1,$\mathbb{R}$) the multiplicative group $\mathbb{R}\backslash\{0\}$ where the easiest choice for $\mu$ is described by $d\mu(a) = \frac{da}{|a|}$. Since $\mathbb{R}$ is the union of the trivial orbit $\{0\}$ and the orbit of 1, the Calderón equation for this special case of the theorem reduces to $\int_{\mathbb{R}\backslash\{0\}} |\widehat{\psi}(\xi)|^2 \frac{d\xi}{|\xi|} = 1$ and the solutions of this equation clearly span a dense subspace of $\mathrm{L}^2(\mathbb{R})$. Since the proof for the general case is identical with that for the special case and, as we observed in 5.1, adapts easily to other situations where we just have a set of dilations rather than a group, every equation of this sort arising from a continuous translation system is called a Calderón equation.

$(iv)$ **Remarks.** Roughly 10 years ago, R. Laugesen, N. Weaver, G. Weiss, and the author of these notes showed that the subgroups $A \subset \mathrm{GL}(n, \mathbb{R})$ for which admissible vectors for $\tau$ exist necessarily have, for $a.e.\ x \in \mathbb{R}^n$, compactness of the stability subgroup $K_x$=$\{a \in A : ax = x\}$, and the semi-direct product G of $A$ and $\mathbb{R}^n$ must be non-unimodular (so the left Haar measure $\nu$ on G is not invariant under all right translations). We also showed that, when G is non-unimodular, a sufficient condition for the existence of admissible vectors is compactness of an $\epsilon$-stabilizer for $a.e.\ x$ (roughly speaking, this means that almost all orbits of $A$ don't return to a small neigborhood of $x$ infinitely often). It's still unknown whether or not compactness of $\epsilon$-stabilizers is necessary for existence of admissible vectors. Recently, Hartmut Führ took a different approach and used a very general measure theoretic argument to show that admissible vectors exist $\Leftrightarrow$ the action of $A$ on $\mathbb{R}^n$ satisfies certain <u>regularity conditions</u> (roughly speaking, this means that, after discarding a Lebesgue null set, the action admits measurable cross-sections containing one point from each orbit).

In brief, the topic of square integrable representations is of great interest in abstract harmonic analysis and it is amusing how one is led into this topic by replacing lattice translations in wavelet and other discrete reproducing function systems by $\mathbb{R}^n$ – translations.

**5.3 The Unified Theorem.** Approximately 10 years ago E. Hernandez, D. Labate, and G. Weiss proved what they called a unified theorem characterizing most Parseval frame lattice-based reproducing function systems by a set (usually countably infinite) of equations, one of which is a Calderón equation. One of the immediate consequences is that a continuous translation system results merely by replacing each lattice with $\mathbb{R}^n$. In this sense, discrete Parseval frame systems are also continuous Parseval frame systems, but the converse of this statement is not true. It makes no sense to try to find all solutions of the Hernandez-Labate-Weiss characterizing equations since there are no available techniques to handle infinitely many complicated equations in infinitely many unknowns. We know lots of tricks for constructing particular sorts of Parseval frame systems (including Gabor systems, ordinary and composite wavelet systems, and more general systems called wave packets). One of the big challenges of the mathematical theory of wavelets is to considerably broaden this "bag of tricks" to obtain many more Parseval frame systems. One way to get started on this is to use perturbations of a Parseval frame system. If one can show that the fact that the original system satisfies the characterizing equations dictates that the perturbations also satisfy these equations, then the result is a family of new Parseval frame systems. Our group used this approach to handle certain perturbed wave packet systems.

**(*ii*) Setting for the Unified Theorem.** Let $I$ be a countable index set, $n$ a fixed member of $\mathbb{N}$, $\{\mathcal{L}_i : i \in I\}$ a countable family of lattices indexed by $I$ and $\{\psi_i : i \in I\}$ a subset of $\mathrm{L}^2(\mathbb{R}^n)\backslash\{0\}$ indexed by $I$. We are then interested in whether or not $\Psi=\{\mathrm{T}_l\psi_i : i \in I\}$ is a Parseval frame for $\mathrm{L}^2(\mathbb{R}^n$, *i.e.* whether or not

$$||f||^2_{L^2(\mathbb{R}^n)} = \sum_{i\in I}\ \sum_{l\in\mathcal{L}_i} |< f,\ \mathrm{T}_l\psi_i >|^2 \qquad (11)$$

holds for each $f \in \mathrm{L}^2(\mathbb{R}^n)$. It's enough to show that (11) holds on a dense subset of $\mathrm{L}^2(\mathbb{R}^n)$, a convenient choice being the collection of functions for which $\widehat{f}$ is continuous and compactly supported with $\widehat{f}(0)$ being non-zero. Furthermore, $||f||^2_{L^2(\mathbb{R}^n)}$ $= ||\mathrm{T}_x f||^2_{L^2(\mathbb{R}^n)}$ for each $x$ so (11) holds $\Leftrightarrow$ the function

$$\omega_f(x) = \sum_{i\in I}\ \sum_{l\in\mathcal{L}_i} |< \mathrm{T}_x f, \mathrm{T}_l\psi_i >|^2 \qquad (12)$$

has the constant value $||f||^2_{L^2(\mathbb{R}^n)}$.

Using bracket functions as in Section 3 and compactness of $f$, each of the terms $\sum_{l\in\mathcal{L}_i} |< \mathrm{T}_x f, \mathrm{T}_l\psi_i >|^2$ in (12) is $\mathcal{L}_i^\perp$ periodic with only finitely many non-zero Fourier coefficients, hence can be written as a trigonometric polynomial
The conclusion below of the Unified Theorem follows <u>provided</u> that it's legitimate, when we sum over $i$, to add up all of the coefficients of $e_\lambda(x)$ for each $\lambda \in \bigcup_{i\in I}(\mathcal{L}_i)^\perp$. This is justified if the sum of the trig polynomials converges absolutely. Hernandez, Labate, and Weiss show show that absolute convergence is guaranteed by a very technical property called the <u>Local Integrability Condition</u> (LIC)

***(iii)* Unified Theorem. (*see* [9])** Suppose $\Psi$ is as in $(ii)$ and satisfies the LIC. Let $\Lambda=\bigcup_{i\in I}(\mathcal{L}_i)^\perp$ and, for each $\lambda \in \Lambda$, let $I_\lambda = \{i \in I : \lambda \in (\mathcal{L}_i)^\perp\}$. Also let $\delta_{\lambda,0}$ be the usual Kronecker $\delta$ function--equal to 1 when $\lambda$=0 and otherwise equal to 0. Then $\Psi$ is a Parseval frame for $\mathrm{L}^2(\mathbb{R}^n) \Leftrightarrow \forall\lambda \in \Lambda$

$$\sum_{i\in I_\lambda} \frac{1}{|\mathcal{L}_i|}\psi_i(\xi)\overline{\psi_i(\xi+\lambda)} = \delta_{\lambda,0} \text{ for } a.e.\ \xi \in (\mathbb{R}^n)\hat{}. \qquad (12)$$

**Note:** For $\lambda$=0, $I_0 = I$ and (12) is then the Calderón equation

$$\sum_{i \in I} \frac{1}{|\mathcal{L}_i|} |\psi_i(\xi)|^2 \quad = 1 \, a.e$$

so, as we mentioned above, at least when the LIC holds, every discrete Parseval frame $\Psi$ as in the Theorem converts to a continuous translation reproducing function system by rescaling.

**Remarks:** In all of the cases where $\Psi$ can be shown directly to be a Parseval frame by other techniques, $e.g.$, Gabor systems and wavelet systems where $\psi_i = D_{a_i}\psi$ with $A = \{a_i : i \in I\}$ a discrete group acting "nicely" on $\mathbb{R}^n$, it turns out that the LIC is automatically satisfied so the characterizing equations hold. There are no examples known where the chararacterizing equations fail to hold and one still has a Parseval frame. Hence, it's unknown whether some other proof could somehow establish validity of the characterizing equations without needing to assume the LIC.

## 6. Conclusion

Although construction of the Haar wavelet system dates back to the early part of the 20th century and construction of the Shannon system was achieved in the mid 1940s, modern wavelet theory essentially dates back only to the work of I. Daubechies, A. Grossman, S. Mallat, and Y. Meyer in the late 1980s. Since then, a very large literature on wavelets has developed. These notes have discussed only the fundamental ideas in wavelet theory and have overlooked many of the deeper ideas discussed in the literature. As we've mentioned, there are many open research questions on higher dimensional wavelets and it's therefore premature to speculate on the "final form" of the higher dimensional theory. Even more speculative is a theory of wavelets on manifolds, *e.g.* on the unit sphere in $\mathbb{R}^n$. Several authors have constructed very interesting wavelets on certain special manifolds. The extent to which their work and that of others will lead to a satisfactory manifold wavelet theory is, to say the least, unclear. Suffice it to say that continued development of wavelet theory will keep a number of researchers very busy for quite some time.

## SHORT WAVELET BIBLIOGRAPHY (for more references, see the bibliographies of the books and papers below plus go to the websites of authors mentioned in the above notes or in other sources on wavelets)

### Books

[1] Brigitte Forster, Peter Massopust, Editors, Four Short Courses on Harmonic Analysis: Wavelets, Frames, Time-Frequency Methods, and Applications to Signal and Image Analysis, Birkhäuser Boston, 2010.

[2] Karlheinz Gröchenig, Aspects of Gabor analysis on locally compact abelian groups, Chapter 6 in Gabor Analysis and Algorithms, edited by H.G. Feichtinger and E. Strömer, Birkhäuser, 2001, pp. 211-231.

[3] Karlheinz Gröchenig, Foundations of Time-Frequency Analysis, Springer Verlag, 2001.

[4] E. Hernandez, G. Weiss, A First Course on Wavelets, Studies in Advanced Mathematics, CRC Press, Boca Raton, FL, 1996

### Articles

[1] D. Bakic, I. Krishtal, E. Wilson, Parseval frame wavelets with dilations by integer matrices of determinant 2 , ACHA 19, 2005, 386-431.

[2] G. Garrigos, E. Hernandez, Hrvoje Sikic, F. Soria, G. Weiss, E. Wilson, Connectivity in the set of Tight Frame Wavelets (TFW) , Glasnik Matematicki 38(58), 2003, 75-98.

[3] P. Gressman, D. Labate, G. Weiss, E. Wilson, Affine, quasi-affine, and co-affine wavelets , published in Beyond Wavelets, edited by J. Stoeckler and G. Welland, Academic Press, 2002, 215-224.

[4] K. Guo, D. Labate, Wang-Q Lim, G. Weiss, E. Wilson, Wavelets with composite dilations and their MRA properties ACHA, 20, 2006, 202-236.

[5] Christopher Heil, Alexander Powell, Gabor Schauder Bases and the Bailian-Low Theorem, J. Math. Physics, **47** (2006), pp.113506-1 through 113506-21.

[6] Eugenio Hernandez, Hrvoje Sikic, Guido Weiss, Edward Wilson, Cyclic Subspaces for Unitary Representations of LCA Groups:Generalized Zak Transforms, accepted for publication in Compositio Mathematica.

[7] Eugenio Hernandez, Hrvoje Sikic, Guido Weiss, Edward Wilson, On the Properties of the Integer Translates of a Square Integrable Function, Contemporary Mathematics, vol. 505, 2010, pp. 233-249.

[8] E. Hernandez, D. Labate, G. Weiss, E.Wilson, Oversampling, quasi-affine frames, and wave packets ACHA, 16, 2004, 111-147.

[9] E. Hernandez, D. Labate, G. Weiss, A unified characterization of reproducing systems generated by a finite family, II, J. Geom. Anal. 12(4), 2002, 615-662.

[10] I. Krishtal, B. Robinson, G. Weiss, E. Wilson, Haar-like wavelets arising from composite dilation systems , J. Geom. Anal., 2007, 87-9.

[11] R. Laugesen, N. Weaver, G. Weiss, E. Wilson, A characterization of the higher dimensional groups associated with continuous wavelets, , J. Geometric Analysis, 12(1), 2001, 89-102.

[12] D. Labate, E. Wilson, Connectivity for Gabor systems , ACHA 17, 2005, 123-136.

[13] D. Labate, G. Weiss, E. Wilson, An approach to the study of wave packet systems , published in Wavelets, Frames, and Operator Theory, C. Heil, P. Jorgensen, and D. Larson (ed.), AMS Contemporary Mathematics Board, 2004, 215-235.

[14] G. Weiss, E. Wilson, The mathematical theory of wavelets , Proceedings of the NATO-ASI Meeting on Harmonic Analysis held in II Ciocco, published in Twentieth
Century Harmonic Analysis--A Celebration, J.S. Byrnes (ed.), Kluwer, 2001, 329-366.