Lecture Notes for Math 648

Professor John Benedetto

University of Maryland, College Park

Fall 2008

Notes Taken by Thomas McCullough[1]

[1]Thanks to Brent Baccala for his proof reading and posting these notes

Comments:

These notes are intended to merely augment and not completely reproduce the recorded lectures. I think it's can be a little tough to see what he's writing, so I think this will be helpful. Professor Benedetto gives significant informative exposition during his lectures (which I appreciate). I have to apologize, but I have largely left this out. This is not out of disrespect, but it's hard to do it much justice without pictures (which I'm not drawing here), and my notes on this are poor (I'm just listening). I think this is much better understood through the videos, anyways.

A note on errors: I'd like these to be professional quality, but I'm producing these notes for my own purposes. I have handwriting like a three year old, but I like to keep my lecture notes. I have absolutely no doubt that this document is riddled with typesetting errors. I welcome your help in finding them. Please email thomas.mccullough@gmail.com, and I'll happily fix them. This will help me, and hopefully you too.

I have tried to stick with his notation. I apologize for the places that I've failed, this is not out of disrespect, just simply my bad habits. I have taken liberty on occasion. In particular, theorems quoted for the result with no proof given I have called propositions, and I have pulled some things I call lemmas out of the proofs to streamline. Any other places that I have taken liberty I explicitly mention. I have cited several things in "Real and Complex Analysis" by Walter Rudin [14], which is the only real analysis book that I have. These things can no doubt be found in any decent analysis book, so consult what you've got.

Contents

Lecture 1

02 September 2008

Let me preface this entire chapter with a reference to [14, Chapter 9]

Fourier Transform Definition and Basic Examples

Definition. Let $f \in L^1(\mathbb{R})$. The *Fourier transform* of f, denoted by $\hat{f}$ is defined as

$$\hat{f}(\gamma) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x) \cdot e^{-2\pi i x\gamma} dx.$$

Note that $f \in L^1(\mathbb{R})$ *does not* imply that $\hat{f} \in L^1(\hat{\mathbb{R}})$ (you'll see an example shortly). In this case, the inverse is tricky to define.

Definition. Define a *dilation* of f by λ as $f_\lambda(t) \stackrel{\text{def}}{=} \lambda \cdot f(\lambda t)$, and consider the effect of this dilation on the Fourier transform. By simple substitution, it is easily demonstrated that

$$\hat{f_\lambda}(\gamma) = \hat{f}(\gamma/\lambda)$$

Example 1.1. Let $f(t) = \mathbb{1}_{[-T,T]}(t)$. Then,

$$\begin{aligned}
\hat{f}(\gamma) &= \int_{\mathbb{R}} \mathbb{1}_{[-T,T]}(t) \cdot e^{-2\pi i t\gamma} dt \\
&= \int_{-T}^{T} e^{-2\pi i t\gamma} dt \\
&= \frac{1}{-2\pi i\gamma} \cdot e^{-2\pi i t\gamma} \Big|_{-T}^{T} \\
&= \frac{1}{\pi\gamma} \cdot \frac{e^{2\pi i T\gamma} - e^{-2\pi i T\gamma}}{2i} \\
&= \frac{\sin(2\pi T\gamma)}{\pi\gamma}
\end{aligned}$$

We define $d(\gamma) = \sin(\gamma)/(\pi\gamma)$, called the *sinc function*. Then, $\hat{f}(\gamma) = d_{2\pi T}(\gamma)$ is a dilation of the sinc function. Note that $\hat{f} \notin L^1(\hat{\mathbb{R}})$.

Example 1.2. Consider $f(t) = e^{-\pi r t^2}$, where $r > 0$. Then,

$$\begin{aligned}
\hat{f}(\gamma) &= \int_{\mathbb{R}} e^{-\pi r t} \cdot e^{-2\pi i t\gamma} dt \\
\frac{d}{d\gamma}\hat{f}(\gamma) &= \frac{d}{d\gamma}\int_{\mathbb{R}} e^{-\pi r t} \cdot e^{-2\pi i t\gamma} dt \\
&= \int_{\mathbb{R}} \frac{d}{d\gamma} e^{-\pi r t} \cdot e^{-2\pi i t\gamma} dt
\end{aligned}$$

This exchange of integral and derivative must be justified, by something like the dominated convergence theorem. Resuming,

$$\begin{aligned}
&= \int_{\mathbb{R}} (-2\pi i) \cdot e^{-\pi r t} \cdot e^{-2\pi i t\gamma} dt \\
&= \frac{i}{r}\int_{\mathbb{R}} \frac{d}{dt}(e^{-\pi r t^2}) \cdot e^{-2\pi i t\gamma} dt \text{ using integration by parts} \\
&= \underbrace{\frac{i}{r} \cdot \left(e^{-\pi r t^2} e^{-2\pi i t\gamma}\right)\Big|_{-\infty}^{\infty}}_{=0} + \frac{-2\pi\gamma}{r}\int_{R} e^{-\pi r t^2} \cdot e^{-2\pi i t\gamma} dt \\
&= \frac{-2\pi\gamma}{r}\hat{f}(\gamma)
\end{aligned}$$

Then, we have the differential equation

$$\frac{d}{d\gamma}\hat{f}(\gamma) = \frac{-2\pi\gamma}{r}\hat{f}(\gamma)$$

We can infer that $\hat{f}(\gamma) = C \cdot e^{g(\gamma)}$, where

$$\frac{d}{d\gamma}g(\gamma) = \frac{-2\pi\gamma}{r} \implies g(\gamma) = \frac{-\pi \cdot \gamma^2}{r}$$

Then, we see that $\hat{f}(\gamma) = C \cdot e^{-\pi \cdot \gamma^2 / r}$ and

$$\begin{aligned}
C = \hat{f}(0) &= \int_{\mathbb{R}} e^{-\pi r t^2} dt \text{ let } u = \pi r t^2 \\
&= \frac{1}{\sqrt{\pi r}}\int_0^{\infty} u^{-1/2} e^{-u} du = \frac{\Gamma(1/2)}{\sqrt{\pi r}} = \frac{1}{\sqrt{r}}
\end{aligned}$$

Then, it follows that $\hat{f}(\gamma) = r^{-1/2} \cdot e^{-\pi\gamma^2/r}$.

Analytic Properties of the Fourier Transform

Proposition 1.3 (Riemann-Lebesgue Lemma)**.** *If $f \in L^1(\mathbb{R})$, then*

$$\lim_{|\gamma| \to \infty} \hat{f}(\gamma) = 0.$$

Proposition 1.4 (Open Problem)**.** *Let f be in the class of all continuous functions which vanish at $\pm\infty$. Is f Fourier equivalent (*same Fourier tranform*) to an $L^1(\mathbb{R})$ function?*

Proposition 1.5. *Suppose that $f \in C^m(\mathbb{R})$ for $m \geq 1$ (differentiable m-times), $f, f', \ldots, f^{(m)} \in L^1(R)$, and*

$$\text{for } k = 0, 1, \ldots, m, \qquad \lim_{x\to\pm\infty} f^{(k)}(x) = 0.$$

Then,

$$\hat{f^{(m)}}(\gamma) = (2\pi i\gamma)^m \cdot \hat{f}(\gamma)$$

This follows from a routine integration by parts. Note the importance of translating the differential operator into a polynomial relation (similar to the Laplace transform).

Definition. Let $f, g \in \mathrm{L}^1(\mathbb{R})$. Recall that the *convolution* of f and g is defined by

$$f * g(t) = \int_{\mathbb{R}} f(t-u) \cdot g(u) du$$

It turns out that the translation invariance of the Lebesgue measure (it is a *Haar* measure) guarantee that convolution is associative and commutative. This is equally valid on any locally compact group using a Haar measure (which is essentially unique).

Proposition 1.6. *If $f, g \in L^1(\mathbb{R})$, then $f * g \in L^1(\mathbb{R})$ and $\widehat{f * g} = \hat{f} \cdot \hat{g}$.*

Proof. This is straight-forward application of Fubini's theorem.

$$\begin{aligned}
\widehat{f * g}(\gamma) &= \int_{\mathbb{R}} f * g(x) \cdot e^{-2\pi i x\gamma} dx \\
&= \int_{\mathbb{R}} \Big(\int_{\mathbb{R}} f(x-y) \cdot g(y)\, dy \Big) \cdot e^{-2\pi i x\gamma}\, dx \\
&= \int_{\mathbb{R}} g(y) \cdot \Big(\int_{\mathbb{R}} f(x-y) \cdot e^{-2\pi i x\gamma}\, dx \Big) dy \\
&= \int_{\mathbb{R}} g(y) \cdot e^{-2\pi i y\gamma} \cdot \Big(\int_{\mathbb{R}} f(x-y) \cdot e^{-2\pi i (x-y)\gamma}\, dx \Big) dy \\
&= \Big(\int_{\mathbb{R}} g(y) \cdot e^{-2\pi i y\gamma}\, dy \Big) \cdot \Big(\int_{\mathbb{R}} f(x) \cdot e^{-2\pi i x\gamma}\, dx \Big) \\
&= \hat{f}(\gamma) \cdot \hat{g}(\gamma)
\end{aligned}$$

□

Note that because convolution is commutative, associative, and distributes with respect to addition, we can consider $\mathrm{L}^1(\mathbb{R})$ as a ring (without unity) with respect to this product. In fact, it is a Banach algebra.

Definition. A sequence (or net) of functions $\{k_n\} \subseteq \mathrm{L}^1(\mathbb{R})$ is called an *approximation to the identity* if the following hold:

1.
$$\forall n, \int k_n(t)dt = 1.$$

2.
$$\exists C > 0 \text{ s.t. } \forall n, \int |k_n(t)|dt \leq C.$$

3.
$$\forall \varepsilon > 0, \lim_{n\to\infty} \int_{|t|>\varepsilon} |k_n(t)|dt = 0.$$

Theorem 1.7. *Let $f \in L^1(\mathbb{R})$ s.t. $\| f \|_{L^1(\mathbb{R})} = 1$. Then the collection $\{f_\lambda\}$ is an approximation to the identity.*

Proof. **Parts (1.) and (2.)** Dilations preserve the value of the respective integrals.
Part (3.) Note that
$$\int_{|t|>\varepsilon} |\lambda f(\lambda t)|dt = \int_{|u|>\varepsilon\cdot\lambda} |f(u)|du$$
By assumption, we have that $f \in \mathrm{L}^1(\mathbb{R})$, so the limit of the RHS must go to 0. □

Theorem 1.8. *Let $f \in L^1(\mathbb{R})$ and $\{k_n\}$ be an approximation to the identity. Then,*
$$\lim_{n\to\infty} \| f - f * k_n \|_{L^1(\mathbb{R})} = 0$$
This motivates the name approximation to the identity, and is a critical idea to establish the fundamental ideas of the Fourier transform.

Proof. Suppose that $\{f_i\} \subseteq C^c(\mathbb{R})$ is a sequence of compactly supported functions in $\mathrm{L}^1(\mathbb{R})$ converging to f (in $\mathrm{L}^1(\mathbb{R})$ norm). It is an elementary property of the $\mathrm{L}^p(\mathbb{R})$ spaces that $C^c(\mathbb{R})$ is dense for $1 \leq p < \infty$. Then, consider that
$$\begin{aligned}\| (f - f * k) - (f_i - f_i * k_n) \|_{\mathrm{L}^1(\mathbb{R})} &= \| (f - f_i) - (f - f_i) * k_n \|_{\mathrm{L}^1(\mathbb{R})} \\ &\leq \| f - f_i \|_{\mathrm{L}^1(\mathbb{R})} + \| (f - f_i) * k_n \|_{\mathrm{L}^1(\mathbb{R})} \text{ Minkowski} \\ &\leq \| f - f_i \|_{\mathrm{L}^1(\mathbb{R})} + \| f - f_i \|_{\mathrm{L}^1(\mathbb{R})} \| k_n \|_{\mathrm{L}^1(\mathbb{R})} \\ &\to 0 \text{ as } i \to \infty\end{aligned}$$
Then, it suffices to assume that $f \in C^c(\mathbb{R})$.

Note that any compactly supported, continuous function is bounded, uniformly continuous, and non-zero on a set of finite measure. Let $M = \| f \|_{\mathrm{L}^\infty(\mathbb{R})}$ and $C = \sup_n \left\{\| f \|_{\mathrm{L}^1(\mathbb{R})}\right\} <$

∞. Given $\varepsilon > 0$, then $\exists \delta > 0$ s.t. $|f(x-y) - f(x)| < \varepsilon/2C, \ \forall |y| < \delta$. Now,

$$\begin{aligned} f(x) - \int_{\mathbb{R}} k_n(y) \cdot f(x-y)dy &= \int_{\mathbb{R}} [k_n(y) \cdot f(x)]dy - \int_{\mathbb{R}} [k_n(y) \cdot f(x-y)]dy \\ &= \int_{\mathbb{R}} k_n(y) \cdot [f(x) - f(x-y)]dy \\ &\leq \int_{\mathbb{R}} |k_n(y)| \cdot |f(x) - f(x-y)|dy \\ &= \overbrace{\int_{|y|\leq\delta} |k_n(y)| \cdot |f(x) - f(x-y)|dy}^{\text{LHS}} + \\ &\qquad \underbrace{\int_{|y|>\delta} |k_n(y)| \cdot |f(x) - f(x-y)|dy}_{\text{RHS}} \end{aligned}$$

Given δ, we know that

$$\text{RHS} \leq 2M \int_{|y|>\delta} |k_n(y)|dy \leq 2M \cdot \frac{\varepsilon}{4M} = \frac{\varepsilon}{2}$$

for sufficiently large n, by property 3 for approximations to the identity.

Also, we know that

$$\text{LHS} \leq \sup_{|y|\leq\delta}\{|f(x) - f(x-y)|\} \int_{\mathbb{R}} |k_n(y)|dy \leq \frac{\varepsilon}{2C} \cdot C = \frac{\varepsilon}{2}.$$

Then, we see that

$$f(x) - \int_{\mathbb{R}} k_n(y) \cdot f(x-y)dy < \varepsilon$$

Then, the desired result follows from the dominated convergence theorem. □

Theorem 1.9. *Suppose that* $H(\gamma) \in L^1(\hat{\mathbb{R}})$ *is a function such that* $0 \leq H(\lambda \cdot \gamma) \leq 1$ *and* $\forall \gamma, \lim_{\lambda\to 0} H(\lambda\cdot\gamma) = 1$ *and* $h_\lambda(x) = \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)\cdot e^{2\pi i x\gamma} d\gamma$ *forms an approximation to the identity*

Then, we have the following:

1. *If* $f \in L^1(\mathbb{R})$*, then*

$$(f * h_\lambda)(x) = \int_{\hat{\mathbb{R}}} H(\lambda \cdot \gamma)\hat{f}(\gamma) \cdot e^{2\pi i x\gamma} d\gamma$$

2. *Suppose that* $f \in L^1(\mathbb{R})$*, and* $\hat{f} \in L^1(\hat{\mathbb{R}})$*. If*

$$g(x) = \int_{\hat{\mathbb{R}}} \hat{f}(\gamma) \cdot e^{2\pi i x\gamma} d\gamma, \ \textit{then} \ \| f - g \|_{L^1(\mathbb{R})} = 0$$

3. If $f \in L^1(\mathbb{R})$ s.t. $\hat{f} = 0$, then $f = 0$.

In [14], he gives the example $H(\gamma) = e^{-|\gamma|}$ and $h_\lambda(t) = 2\lambda/(\lambda^2+x^2)$. Professor Benedetto was using $H(\gamma) = \mathbb{1}_{[-1/2\pi,1/2\pi]}(\lambda) \cdot (1 - 2\pi|\gamma|)$, but the only important properties for the two are stated above.

Proof. **Part (1.)** This is a simple application of Fubini's theorem.

$$\begin{aligned}(f * h_\lambda)(x) &= \int_{\mathbb{R}} f(x-y)dy \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)\cdot e^{2\pi i y\gamma}d\gamma \\ &= \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)d\gamma \int_{\mathbb{R}} f(x-y)\cdot e^{2\pi i y\gamma}dy \\ &= \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)d\gamma \int_{\mathbb{R}} f(y)\cdot e^{2\pi i(x-y)\gamma}dy \\ &= \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)\hat{f}(\gamma)\cdot e^{2\pi i x\gamma}d\gamma\end{aligned}$$

Part (2.) First, note that

$$\| f - g \|_{L^1(\mathbb{R})} \leq \int_{\mathbb{R}} \left| f(x) - (f * h_\lambda)(x) \right| dx + \int_{\mathbb{R}} \left| (f * h_\lambda)(x) - \int_{\hat{\mathbb{R}}} H(\lambda\cdot\gamma)\cdot\hat{f}\cdot e^{2\pi i x\gamma}d\gamma \right| dx,$$

by the triangle inequality. Letting $\lambda \to 0$, we know that the left component of the above goes to 0 by the assumption that $\{h_\lambda\}$ forms an approximation to the identity. Similarly, we assumed that $H(\lambda\cdot\gamma) \to 1$ and the dominated convergence theorem implies the right component goes to 0. Then, we're done.

Part (3.) A direct corollary to the previous part. □

This establishes the basic properties of the Fourier transform. The next results require more sophisticated analytic techniques.

Proposition 1.10 (Plancherel Theorem)**.** *There exists a unique linear isometry*

$$\tilde{F} : L^2(\mathbb{R}) \to L^2(\hat{\mathbb{R}})$$

such that

1. *$\forall f \in L^2(\mathbb{R})$, $\| f \|_{L^2(\mathbb{R})} = \| \tilde{F}(f) \|_{L^2(\hat{\mathbb{R}})}$ (isometry).*
2. *$\forall f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, $\tilde{F}(f) = \hat{f}$ (extends our definition of Fourier transform).*

The proof of this is quite similar to the above, and the technical details are omitted here. It follows from the fact that $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, $\| f \|_{L^2(\mathbb{R})} = \| \hat{f} \|_{L^2(\hat{\mathbb{R}})}$. The rest follows from the fact that $L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, and we have an isometry mapping a dense subset of $L^2(\mathbb{R})$ onto a dense subset of $L^2(\hat{\mathbb{R}})$. The completeness of these two spaces guarantee the existence and uniqueness of $\tilde{F}$.

Proposition 1.11 (Parseval's Formula)**.** *Let $f, g \in L^2(\mathbb{R})$, then*

$$\int_{\mathbb{R}} f(x)\cdot\overline{g(x)}\,dx = \int_{\hat{\mathbb{R}}} \hat{f}(\gamma)\cdot\overline{\hat{g}(\gamma)}\,d\gamma$$

Fourier Series

Definition. Let $\Omega > 0$ and $F\colon \mathbb{R} \to \mathbb{C}$ be 2Ω periodic. Define $\mathbb{T}_{2\Omega} = \mathbb{R}/(2\Omega)$, and suppose that

$$\int_K |F|\, dx < \infty \text{ for } \forall K \subset \mathbb{R} \text{ of finite measure.}$$

Then, we say that $F \in \mathrm{L}^1(\mathbb{T}_{2\Omega})$. The *Fourier Series* of such an F is denoted

$$S(F)(\gamma) = \sum_{n\in\mathbb{Z}} f[n] \cdot e^{-2\pi i n\gamma/2\Omega},\ f[n] \overset{\text{def}}{=} \frac{1}{2\Omega} \cdot \int_{-\Omega}^{\Omega} F(\gamma) \cdot e^{2\pi i n\gamma/2\Omega} d\gamma$$

Note that there is no guarantee that this series converges, and sometimes it doesn't. There Riemann-Lesbesgue lemma still holds in this setting.

Proposition 1.12 (Dirichlét). *Let $F \in L^1(\mathbb{T}_{2\Omega})$, and F be differentiable at γ_0. Then, $S(F)(\gamma_0) = F(\gamma_0)$, in the sense that*

$$\lim_{M,N\to\infty} \sum_{n=-M}^{N} f[n] \cdot e^{-2\pi i n\gamma/2\Omega} = F(\gamma_0)$$

Proposition 1.13. *Let $F \in L^1(\mathbb{T}_{2\Omega})$. then, $\forall[\alpha, \beta] \subset [-\Omega, \Omega]$ we have*

$$\sum_{n\in\mathbb{Z}} f[n] \cdot \int_{\alpha}^{\beta} e^{-2\pi i n\gamma/2\Omega} d\gamma = \int_{\alpha}^{\beta} F(\gamma) d\gamma$$

This is true, regardless of the convergence of the Fourier series.

Periodization of an $\mathbf{L}^1(\mathbb{R})$ Function

Definition. Let $T > 0$ and $f \in \mathrm{L}^1(\mathbb{R})$. The *$T$-periodization* of f is the T-periodic function

$$\mathring{f}_T(t) = \sum_{n\in\mathbb{Z}} f(t - nT)$$

Note that

$$\begin{aligned}
\int_0^T |\mathring{f}(t)|\, dt &= \int_0^T \Big|\sum_{n\in\mathbb{Z}} f(t - nT)\Big|\, dt \\
&\le \sum_{n\in\mathbb{Z}} \int_0^T |f(t - nT)|\, dt \text{ (triangle inequality and Fubini)} \\
&= \sum_{n\in\mathbb{Z}} \int_{-nT}^{(1-n)T} |f(u)|\, du \\
&= \int_{\mathbb{R}} |f(u)|\, du = \|f\|_{\mathrm{L}^1(\mathbb{R})},
\end{aligned}$$

which establishes that $\mathring{f}_T \in \mathrm{L}^1(\mathbb{T}_T)$ and that the expression defining $\mathring{f}_T$ converges a.e.

Theorem 1.14. *Let $f \in L^1(\mathbb{R})$ and $T > 0$, and $\{\mathring{f}_T[n]\}_{n\in\mathbb{Z}}$ be the sequence of Fourier series coefficients for $\mathring{f}_T$. Then,*

$$\mathring{f}_T[n] = \frac{1}{T} \cdot \hat{f}\Big(\frac{-n}{T}\Big)$$

Proof.

$$\begin{aligned}
\mathring{f}_T[n] &= \frac{1}{T}\int_0^T \mathring{f}_T(t) \cdot e^{2\pi int/T}\, dt \\
&= \sum_{m\in\mathbb{Z}} \frac{1}{T}\int_0^T f(t-mT) \cdot e^{2\pi int/T}\, dt \text{ (Fubini)} \\
&= \sum_{m\in\mathbb{Z}} \frac{1}{T}\int_{-mT}^{(1-m)T} f(u) \cdot e^{2\pi in(u+mT)/T}\, du \\
&= \frac{1}{T}\sum_{m\in\mathbb{Z}} \int_{-mT}^{(1-m)T} f(u) \cdot e^{2\pi inu/T}\, du \\
&= \frac{1}{T}\int_{\mathbb{R}} f(u) \cdot e^{2\pi inu/T}\, du \\
&= \frac{1}{T} \cdot \hat{f}\Big(\frac{-n}{T}\Big)
\end{aligned}$$

□

Corollary 1.15 (Poisson Summation Formula)**.** *Suppose that $f \in L^1(\mathbb{R})$. Suppose that $S(\mathring{f}_T)(t) = \mathring{f}_T(t)$ for some $t \in \mathbb{T}_T$ (i.e. if $\mathring{f}_T$ is smooth at t). Then,*

$$T\sum_{n\in\mathbb{Z}} f(t+nT) = \sum_{n\in\mathbb{Z}} \hat{f}\Big(\frac{-n}{T}\Big) \cdot e^{2\pi int/T}$$

Proof. By assumption,

$$T\sum_{n\in\mathbb{Z}} f(t+nT) = T \cdot S(\mathring{f}_T)(t) = T \cdot \mathring{f}_T(t) = T\sum_{n\in\mathbb{Z}} \mathring{f}_T[n] \cdot e^{-2\pi int/T} = \sum_{n\in\mathbb{Z}} \hat{f}\Big(\frac{-n}{T}\Big) \cdot e^{-2\pi int/T}$$

□

Lecture 2

09 September 2008

Proposition 2.1 (Poisson Summation Formula)**.**

$$T \cdot \sum_{n \in \mathbb{Z}} f(t+nT) = \sum_{n \in \mathbb{Z}} \hat{f}(n/T) e^{2\pi i n t/T}$$

Proof. Proof omitted here. See [2, pgs. 505-523] for a proof. □

Example 2.2. It is the case that $\exists f \in \mathrm{L}^1(\mathbb{R})$, f continuous, $f(n) = 0,\ \forall n \in \mathbb{Z}$, $\hat{f}(n) = 0\ \forall n \neq 0$ and $\hat{f}(0) = 1$.

Proof. No proof of this claim will be provided. This example is intended to illustrate that the statement in Proposition 2.1 doesn't hold universally. The conditions under which it does hold have remained unstated, but there clearly are some. □

For clear notation in what follows, define

$$\mathrm{PW}_\Omega \stackrel{\text{def}}{=} \{f \in \mathrm{L}^2(\mathbb{R}) \mid \operatorname{supp}(\hat{f}) \subset [-\Omega, \Omega]\} \text{ and } \tau_u s(t) \stackrel{\text{def}}{=} s(t-u)$$

where PW stands for Paley-Weiner.

Theorem 2.3 (Classical Sampling Theorem)**.** *Let T, Ω be positive real numbers such that $0 < 2T\Omega \leq 1$. Let $s \in \mathrm{PW}_{1/2T}$ with $\hat{s}(\gamma) = 1$ for $\gamma \in [-\Omega, \Omega]$, $\hat{s}(\gamma) = 0$ for $\gamma \notin [-1/2T, 1/2T]$, and $\hat{s} \in L^\infty(\hat{R})$. Then,*

$$\forall f \in \mathrm{PW}_\Omega,\ f = T \cdot \sum_{n \in \mathbb{Z}} f(nT) \tau_{nT} s$$

in the L^2 norm.

Proof. Note that by the Plancherel Theorem [14, p. 186], we have:

$$\begin{aligned} \| f - T \cdot \sum_{|n| \leq N} f(nT) \tau_{nT} s \|_{\mathrm{L}^2(\mathbb{R})} &= \| \hat{f}(\gamma) - T \cdot \sum_{|n| \leq N} f(nT) \cdot e^{-2\pi i n T \gamma} \cdot \hat{s}(\gamma) \|_{\mathrm{L}^2(\hat{\mathbb{R}})} \\ &= \| \hat{f}(\gamma) - T \cdot \sum_{|n| \leq N} f(nT) \cdot e^{-2\pi i n T \gamma} \cdot \hat{s}(\gamma) \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})} \end{aligned} \tag{2.1}$$

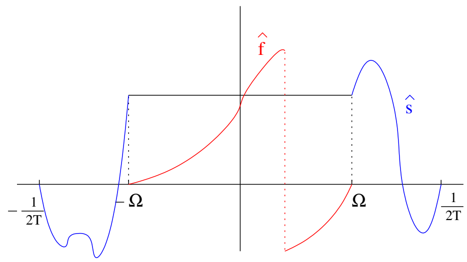


Figure 2.1: Example f and s

Define $G \in \mathrm{L}^2(\mathbb{T}_{1/T})$, where $\mathbb{T}_{1/T} \stackrel{\text{def}}{=} \mathbb{R}/((1/T)\mathbb{R})$, by

$$G(\gamma) = \begin{cases} \hat{f}(\gamma), & |\gamma| < \Omega \\ 0, & \Omega \leq \gamma < 1/2T \end{cases}$$

Then, by [14, p. 186], the Fourier series of G is given by

$$\sum_{|n| \leq N} \check{G}[n] \cdot e^{-2\pi i n T \gamma}$$

where

$$\check{G} = \int_{\mathbb{T}_{1/T}} G(\gamma) \cdot e^{2\pi i n T \gamma} d\gamma = T \int_{-\Omega}^{\Omega} \hat{f}(\gamma) e^{2\pi i n T \gamma} d\gamma = T f(nT)$$

Then, continuing our calculation

$$\begin{aligned}(2.1) &= \| \hat{f} - \sum_{|n| \leq N} \check{G}(n) \cdot e^{-2\pi i n T \gamma} \cdot \hat{s} \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})} \\ &= \| \hat{f} - G \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})} + \| G - S_N(G) \cdot \hat{s} \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})}\end{aligned}$$

The first term goes to zero, again by [14, p. 186], so we are left with

$$\begin{aligned}\| G - S_N(G) \cdot \hat{s} \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})} &= \| \hat{s}(G - S_N(G)) \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})} \\ &\leq \| \hat{s} \|_{\mathrm{L}^\infty(\hat{\mathbb{R}})} \cdot \| G - S_N(G) \|_{\mathrm{L}^2(\widehat{[-\frac{1}{2T}, \frac{1}{2T}]})}\end{aligned}$$

where the last inequality is Hölder's. This last term again goes to 0 by [14, p. 186]. That completes the proof. □

Example 2.4. Let $2T\Omega = 1$ and $s_\Omega(t) = d_{2\pi\Omega}(t) = \sin(2\pi\Omega t)/\pi t$. Recall that the inverse Fourier transform of $[-\Omega, \Omega]$ is s_Ω.

Example 2.5. Set $\varphi(t) = s_\Omega(t)/\sqrt{2\Omega}$. Let $V_0 = \overline{\text{span}}\{\tau_{nT}\varphi\}$. This leads directly to an MRA, which will be discussed later. Let $\psi(t) = (1/\sqrt{2\Omega}) \cdot (s_{2\Omega}(t) - s_\Omega(t))$. This is known as the *Shannon dyadic wavelet*.

Now, for some applications of PSF (2.1):

1. $T\sum \delta_{nT} = \sum e^{2\pi i t n/T}$ - engineering notation for dirac delta function (distribution)
2. Classical Sampling Theorem and relations to Locally Compact Abelian Groups
3. Euler-MacLaurin Formula: $T \cdot \sum_0^\infty f(nT) = \int_0^\infty f(t)dt$ + error terms
4. Jacobi Formula: $\vartheta(t) = \sum e^{-\pi n^2 t}$.
   (a) $\forall t > 0, \vartheta(t) = \frac{1}{\sqrt{t}} \cdot \vartheta(\frac{1}{t})$
   (b) Diffusion Equations
   (c) Statistical Mechanics
   (d) Automorphic forms & Elliptic functions
   (e) Deligne's proof of the Ramanajan conjecture
   (f) Selberg trace formula is CST in number theoretic, non-abelian setting.

Given $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$. Define the Fourier transform of $f : \mathbb{Z}_n \to \mathbb{C}$ as $F : \mathbb{Z}_n \to \mathbb{C}$ where

$$F[n] = \sum_{m \in \mathbb{Z}_n} f(m) e^{-2\pi i m n/N}$$

Theorem 2.6 (Inversion Formula)**.** *Given f and F as stated, then*

$$f(m) = \frac{1}{N} \sum_{n \in \mathbb{Z}_n} F(n) e^{2\pi i m n/N}$$

Proof. This shakes out immediately from the fact that

$$\sum_{n=0}^{N} e^{2\pi i n/N} = 0.$$

□

Theorem 2.7 (Discrete Fourier Transform)**.** *Let $\Omega > 0$, $N \in 2\mathbb{N}$, and T s.t. $2\Omega T = 1$. If $f \in \mathrm{PW}_\Omega$, then consider the dilation f_T as a function $f_T : \mathbb{Z} \to \mathbb{C}$ (in addition to being a continuous complex valued function on $\mathbb{R}$), by $m \mapsto f_T[m]$. Define $W_N = e^{2\pi i/N}$ for notational purposes. Assume that $f_T \in \ell^1(\mathbb{Z})$ and suppose that $\hat{f}$ is continuous on $[-\Omega, \Omega]$. Then, $\forall n \in (-N/2, N/2)$, we have*

$$\hat{f}\left(\frac{2\Omega n}{N}\right) = \hat{f}\left(\frac{n}{NT}\right) = \sum_{m=0}^{N-1} (f_T)^o[m] \cdot W_N^{mn}, \text{ where } (f_T)^o_N = T \sum_{k \in \mathbb{Z}} f((m+kN) \cdot T)$$

Proof. By the CST (2.3), we have that

$$f = T \sum f(mT) \cdot \tau_{mT} d_{2\pi\Omega} \implies \hat{f} = T \sum f(mT) \cdot e_{-mT} \cdot \mathbb{1}_{[-\Omega,\Omega]}$$

where $e_r(\gamma) = e^{2\pi i r\gamma}$.

If $n \in (-N/2, N/2)$, then

$$\begin{aligned}
\hat{f}\left(\frac{2\Omega n}{N}\right) &= T \sum f(mT) \cdot e^{-2\pi i mT2\Omega n/N} \qquad (2T\Omega = 1) \\
&= T \sum_m \sum_{p=mN}^{mN+N-1} f(pT)e^{-2\pi i pn/N} \\
&= T \sum_m \sum_{j=0}^{N-1} f((j + mN)T) \cdot e^{-2\pi i(jm/N+mN)}
\end{aligned}$$

Rearranging the sum completes the proof. □

Now, for some historical motivation for wavelets.

Definition. Let $g \in L^2(\mathbb{R})$, and $a, b > 0$. The *Gabor or Weyl-Heisenberg system of Coherent States* is the sequence $\{g_{m,n} \mid (m, n) \in \mathbb{Z} \times \mathbb{Z}\}$. Where

$$g_{m,n}(t) = e^{2\pi i tmb} \cdot g(t - na) = e_{mb}(t) \cdot \tau_{na} g(t).$$

Note that $\hat{g}_{m,n}(\gamma) = \tau_{mb}(e_{-na} \cdot \hat{g})(\gamma)$. This arises as a tool in Quantum Mechanics.

Definition. Let $\psi \in L^2(\mathbb{R})$. The *Dyadic Wavelet or Affine System* for ψ is the sequence $\{\psi_{m,n} \mid (m, n) \in \mathbb{Z} \times \mathbb{Z}\}$, where

$$\psi_{m,n}(t) = 2^{m/2} \cdot \psi(2^m \cdot t - n) \implies \hat{\psi}_{m,n}(\gamma) = 2^{-m/2} \cdot e_{-n} \cdot \hat{\psi}(\gamma \cdot 2^{-m})$$

This arises in conjunction with the so called *Affine Group*, the group of affine transformations of $\mathbb{R}$.

Some wavelets references: for a mathematical treatment see [10], for an applied math treatment see [3], for an engineering treatment see [9].

Wavelets were developed independently in numerous disparate fields, from distinct efforts and without significant cross fertilization until relatively recently.

In mathematics, this work was motivated by work in algebraic bases for function spaces, the study of Fourier tranforms/series, and splines. Significant work was performed by Haar (1909) in his PhD thesis, Franklin (1927), and Stromburg (1970's), Littlewood-Paley theory, and the Calderón formula.

In physics, the work was motivated by the above Gabor systems, in the work of Von Neumann (1920's or 1930's), Heisenberg, and Weyl.

In engineering, the work was motivated by STFT (Short Time Fourier Transform), speech processing (1970's), two aspects of multi-resolution analysis - Quadratic Mirror Filters (1970's) and Image Processing (pyramidal schemes), the radar-ambiguity function (1953), and Walsh functions (primordial wavelet packets).

Proposition 2.8 (Alberto Calderón). *$\exists\psi$ such that $\forall f \in L^2(\mathbb{R})$*

$$f(t) = \int_{\mathbb{R}} \psi_{1/u} * \psi_{1/u} * f \frac{du}{u}$$

Proof. For a proof, see [1, 2.2.2 (c.)]. It will not be proved here. To see what ψ must be like, let's take the Fourier transform. Note that

$$\hat{f}(\gamma) = \hat{f}(\gamma) \cdot \int_{\mathbb{R}} [\hat{\psi}(u\gamma)]^2 \frac{du}{u} \implies \int_{\mathbb{R}} [\hat{\psi}(u\gamma)]^2 \frac{du}{u} = 1 \text{ (almost everywhere)}$$

The establishes the continuous wavelet transform. □

Proposition 2.9 (Ingrid Daubechies). *Given $r \geq 1$, then $\exists\psi \in C_c^{(r)}(\mathbb{R})$ such that $\{\psi_{m,n}\}$ is an orthonormal basis for $L^2(\mathbb{R})$.*

Proof. No proof given. This establishes a wavelet basis of arbitrary smoothness for $\mathrm{L}^2(\mathbb{R})$. This is an important extension to Haar's work, which established a non-smooth (step function, in fact) wavelet basis for $\mathrm{L}^2(\mathbb{R})$. □

Lecture 3

16 September 2008

The next step in the evolution of wavelet theory was *Multi-Resolution Analysis (MRA)*.

Definition. (Intuitive) Set

$$f_M = \sum_{m \leq M} \sum_{n \in \mathbb{Z}} \langle f, \psi_{m,n} \rangle \cdot \psi_{m,n}$$

Assume that ψ satisfies $\operatorname{supp} \psi \subseteq [-1/2, 1/2]$. Then, $\operatorname{supp} \psi_{m,n} \subseteq I_{m,n} \overset{\text{def}}{=} [n2^{-m} - 2^{-(m+1)}, n2^{-m} + 2^{-(m+1)}]$. Note that

$$f_{M+1} = f_M + \langle f, \psi_{M+1,n}, \cdot \rangle \psi_{M+1,n}$$

In other words, f_{M+1} *deblurs* f_M by adding details at a finer scale, on intervals of length $2^{-(m+2)}$. This is the essence of *MRA*.

Definition. (Formal) This was probably first formalized by Y. Meyer. The pair $\{V_j\}_{j \in \mathbb{Z}}, \varphi$ is an *MRA* of $L^2(\mathbb{R})$ if

1. each V_j is a closed subspace (of $L^2(\mathbb{R})$.
2. $V_j \subseteq V_{j+1}$ for each $j \in \mathbb{Z}$.
3. $\bigcup V_j = L^2(\mathbb{R})$ and $\bigcap V_j = 0$ (the zero function).
4. $f(t) \in V_j \iff f(2t) \in V_{j+1}$.
5. $f \in V_0 \iff \forall k \in \mathbb{Z}, \tau_k f \in V_0$.
6. $\varphi \in V_0$ and $\{\tau_k \varphi\}_{k \in \mathbb{Z}}$ for an orthonormal basis (ONB) for V_0.

For example, $\varphi = \mathbb{1}_{[0,1)}$ and $V_0 = \overline{\operatorname{span}}\{\tau_k \varphi\}$, this is the Haar system.

Proposition 3.1. *Given an MRA $\{V_j\}, \varphi$ of $L^2(\mathbb{R})$, then there exists an* explicitly constructible ψ *s.t.* $\{\psi_{m,n}\}$ *is an ONB for $L^2(\mathbb{R})$.*

Outline of Meyer's Algorithm: W_j is defined s.t. $V_j \oplus W_j = V_{j+1}$.

1. $\exists \psi \in W_0$ s.t. $\{\tau_k \psi\}$ is an ONB for W_0 and $\{\psi_{m,n}\}$ form an ONB for $L^2(\mathbb{R})$.
2. $\exists h_0[n]$ s.t.

$$\varphi(t) = \sqrt{2} \cdot \sum h_0[n] \cdot \varphi(2t - n) \implies \sqrt{2} \cdot \hat{\varphi}(2\gamma) = H_0(\gamma) \cdot \hat{\varphi}(\gamma).$$

Where $H_0(\gamma) = \sum h_0[n] \cdot e^{-2\pi i n \gamma}$. Then,

$$\psi(t) = \sqrt{2} \cdot \sum h_1[n] \cdot \varphi(2t - n) \implies \sqrt{2} \cdot \hat{\psi}(2\gamma) = -e^{-2\pi i \gamma} \cdot \overline{H_0}(\gamma + 1/2) \cdot \hat{\varphi}(\gamma)$$

where $h_1[n] = (-1)^n \cdot \overline{h_0}[-n + 1]$. So, $|H_0(\gamma)|^2 + |H_0(\gamma + 1/2)|^2 = 2$.

Proposition 3.2 (Heisenberg Uncertainty Principle)**.** *If $f \in L^2(\mathbb{R})$, $t_0 \in \mathbb{R}$, $\gamma_0 \in \hat{\mathbb{R}}$, then*

$$\| f \|^2_{L^2(\mathbb{R})} \leq 4\pi \cdot \| (t - t_0) \cdot f(t) \|_{L^2(\mathbb{R})} \cdot \| (\gamma - \gamma_0) \cdot \hat{f}(\gamma) \|_{L^2(\hat{R})}$$

There are a series of Fourier uncertainty principles spawned by this idea.

This gives rise to two extreme cases:

- Suppose that $f \in L^2_{\text{loc}}(\mathbb{R})$ so that f is $L^2(I)$ for any finite interval I, and $a \in [-3/2, -1/2)$. Suppose that $f(t)$ is asymptotic to $|t|^a$. then, $f \in L^2(\mathbb{R})$ and $\int |t|^2 \cdot |f(t)|^2 dt = \infty$.
- Suppose that $f = \mathbb{1}_{[-T,T]}$ so that $\hat{f} = d_{2\pi T}$ and $\int |\gamma|^2 \cdot |\hat{f}(\gamma)|^2 d\gamma = \infty$.

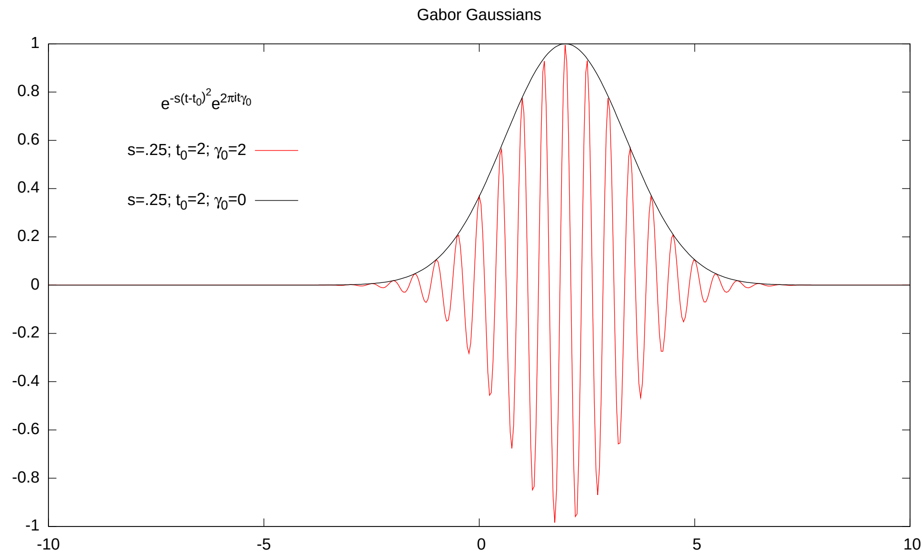
Example 3.3. Suppose that

$$g(t) = \sqrt{\frac{2s}{\pi}} \cdot e^{-s(t-t_0)^2} \cdot e^{2\pi i t \gamma_0} \implies$$
$$\sigma^2 t = 4\pi \| (t - t_0) \cdot g(t) \|^2_{L^2(\mathbb{R})} = \pi / s$$
$$\sigma^2 s = 4\pi \| (\gamma - \gamma_0) \cdot \hat{g}(\gamma) \|^2_{L^2(\hat{\mathbb{R}})} = s/\pi.$$

In this example, we get equality in the uncertainly principle. Apparently, this is an if and only if situation. If we view the uncertainty principle as a product of variances, then this situation for g seems to maximize the effectiveness of the process. This approach was an important part of Gabor's contribution - "the best utilization of the information area". He wanted to decompose every signal into a collection of gaussians, to minimize the uncertantity. That is, write

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

As it turns out, this doesn't quite work. It was a significant contribution to the field, because it stimulated a lot of ideas.



Proposition 3.4 (Balian - Low Theorem)**.** *Given $a, b > 0$ with $ab = 1$. Let $f \in L^2(\mathbb{R})$, and $f_{m,n}(t) = e^{2\pi imbt} \cdot f(t - na)$. If $\{f_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$, then*

$$\int_{\mathbb{R}} |t \cdot f(t)|^2 dt = \infty \text{ or } \int_{\hat{\mathbb{R}}} |\gamma \cdot \hat{f}(\gamma)|^2 d\gamma = \infty$$

It is clear that Gabor's idea is incompatible with this fact.

Example 3.5 (Morlet Wavelet)**.** Let $\psi(t) = e^{-\pi \cdot t^2} \cdot (e^{2\pi it\gamma_0} - e^{-\pi\gamma_0^2})$ and $\psi_{m,n}(t) = 2^{m/2} \cdot \psi(2^m \cdot t - n)$ be the dyadic system of dialates and translates.

The real idea is the "same number of cycles for low, medium, and high frequencies". See [12].

Proposition 3.6 (Gabor Decomposition)**.** *This notation is a bit stiff. Suppose that $T, \Omega > 0$, $2T\Omega \leq 1$, and $g \in \mathrm{PW}\, 1/2T$ with $\hat{g} \in L^\infty(\hat{\mathbb{R}})$, $\hat{g} = 1$ on $[-\Omega, \Omega]$. If $2T\Omega < 1$, then there are other conditions incompletely stated... ($\hat{g}$ continuous, $\hat{g} > 0$ on $[-1/2T, 1/2T]$, maybe more?). Set*

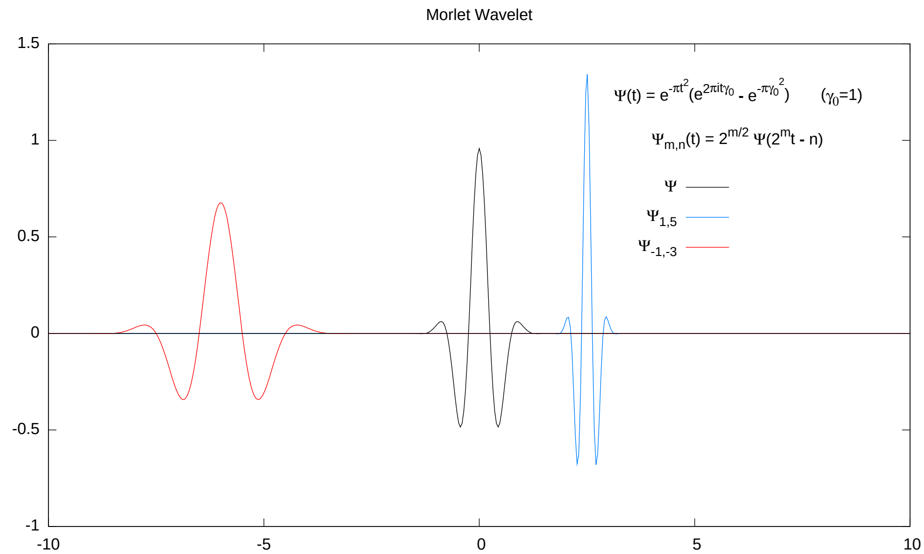
$$G(\gamma) = \sum |\hat{g}(\gamma - mb)|^2 \quad s(t) = (\hat{g}/G)\check{}(t)$$

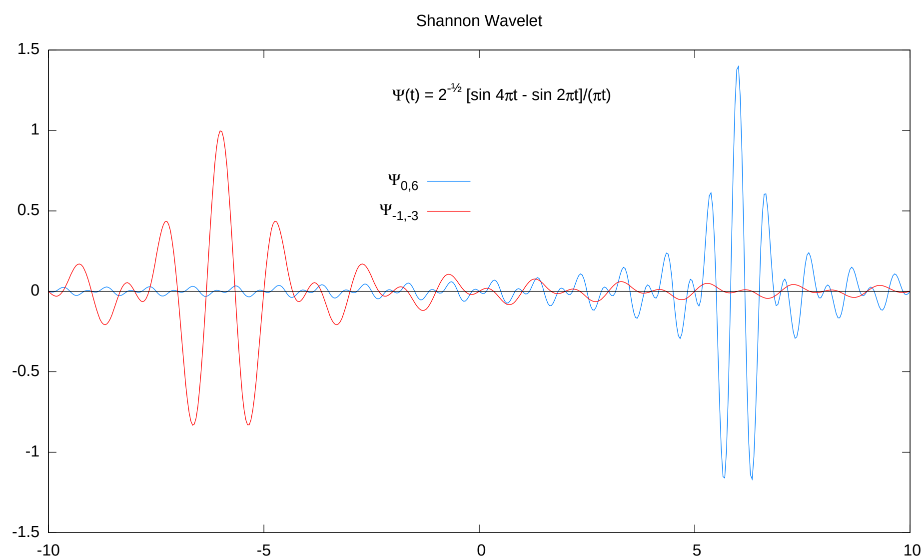
Then, $\forall f \in L^2(\mathbb{R})$,

$$f = T \cdot \sum \langle \hat{f}, e_{nt} \tau_{mb} \hat{g} \rangle \cdot \tau_{-mT}(e_{mb} s)$$

Then, by theorem 2.3,

$$\forall f \in \mathrm{PW}_\Omega, f = T \sum f(nT) \tau_{mT} s$$





Theorem 3.7 (Shannon Wavelet Decomposition)**.** *Let* $\Omega > 0$*,*

$$\varphi = (1/\sqrt{2\Omega}) \cdot d_{2\pi\Omega}$$

$$\psi = (1/\sqrt{2\Omega}) \cdot (d_{2\pi(2\Omega)} - d_{2\pi\Omega}) \text{ so}$$

$$\sqrt{2\Omega} \cdot \hat{\psi} = \mathbb{1}_{[-2\Omega,-\Omega)} + \mathbb{1}_{(\Omega,2\Omega]}$$

$$\sqrt{2\Omega} \cdot \hat{\psi}_{m,0}(\gamma) = 2^{-m/2} \cdot \sqrt{2\Omega} \cdot \hat{\psi}\left(\frac{\gamma}{2^m}\right)$$

Let $f \in L^2(\mathbb{R})$, $\hat{f} = F$ (notational convenience), $\Omega > 0$, φ and ψ as above. Then,

$$f = \sqrt{2\Omega} \cdot f * \varphi + \sum_{m=0}^{\infty} \sum_{n \in \mathbb{Z}} d_{m,n} \psi_{m,n/(4\Omega)} \implies$$
$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} d_{m,n} \cdot \psi_{m,n/(4\Omega)} \text{ in the } L^2(\mathbb{R}) \text{ norm.}$$
$$d_{m,n} = \frac{1}{\sqrt{2\Omega} \cdot 2^{(m/2+1)}} \cdot \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} F(\gamma) \cdot \left(\mathbb{1}_{[-2^m\Omega, -2^m\Omega)} + \mathbb{1}_{(2^m\Omega, 2^{m+1}\Omega]} \right)(\gamma) \cdot e^{2\pi i n\gamma/(2^{m+2}\Omega)} d\gamma$$

Proof.

$$F(\gamma) = \sqrt{2\Omega} \cdot \left(F(\gamma) \cdot \hat{\varphi}(\gamma) + \sum_{m=0}^{\infty} F(\gamma) \cdot \hat{\psi}\left(\frac{\gamma}{2^m}\right) \right) = \sqrt{2\Omega} \cdot \sum_{m \in \mathbb{Z}} F(\gamma) \cdot \hat{\psi}\left(\frac{\gamma}{2^m}\right).$$

Set $F_m(\gamma) = \sqrt{2\Omega} \cdot F(\gamma) \cdot \hat{\psi}(\gamma/2^m)$, and $f_m = \check{F}_m$. Then, $\forall m \in \mathbb{Z}$, $f_m \in \mathrm{PW}_{2^{m+1}\Omega}$ and $\operatorname{supp} F_m \subseteq [-2^{m+1}\Omega, -2^m\Omega) \cup (2^m\Omega, 2^{m+1}\Omega]$.

Then, we can consider F_m as a $2^{m+2}\Omega$ periodic function on $\hat{\mathbb{R}}$ with

$$\sqrt{2\Omega} \cdot F(\gamma) \cdot \hat{\psi}\left(\frac{\gamma}{2^m}\right) \text{ on } [-2^{m+1}\Omega, 2^{m+1}\Omega]$$
$$S(F_m)(\gamma) = \sum_{m \in \mathbb{Z}} c_{m,n} \cdot e^{-2\pi i m\gamma/(2^{m+2}\Omega)} \text{ (Fourier series)}$$

From the basic facts of Fourier series, we know that

$$\begin{aligned} f_m(t) &= \int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} F_m(\gamma) e^{2\pi i t\gamma} d\gamma \\ &= \sqrt{2\Omega} \cdot \sum_{m \in \mathbb{Z}} c_{m,n} \cdot \underbrace{\int_{-2^{m+1}\Omega}^{2^{m+1}\Omega} \hat{\psi}\left(\frac{\gamma}{2^m}\right) \cdot e^{2\pi i (t - m/(2^{m+2}\Omega))\gamma} d\gamma}_{\psi_{m,n/(4\Omega)}} \text{ in } \mathrm{L}^2(\mathbb{R}) \text{ norm.} \end{aligned}$$

But, remember that

$$f = \sum_{m \in \mathbb{Z}} f_m = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \sqrt{2\Omega} \cdot c_{m,n} \cdot \psi_{m,n/(4\Omega)}$$

□

Note:

- $\{\psi_{m,n/(4\Omega)}\}$ is not orthogonal in $\mathrm{L}^2(\mathbb{R})$, even though

$$\left\{ \frac{1}{2^{m/2} \cdot \sqrt{2\Omega}} \cdot e^{-2\pi i (n/(2^{m+2}\Omega))\gamma} \right\}$$

  is an ONB for $\mathrm{L}^2(\mathbb{T}_{2^{m+2}\Omega})$.

- $\{\psi_{m,n/(2\Omega)}\}$ is an orthonormal sequence in $L^2(\mathbb{R})$. Therefore, if we show

$$\forall f \in L^2(\mathbb{R}), \quad \sum_{m,n\in\mathbb{Z}} |\langle f, \psi_{m,n/(2\Omega)}\rangle|^2 = \|f\|^2_{L^2(\mathbb{R})},$$

  then we can conclude that $\{\psi_{m,n/(2\Omega)}\}$ is an ONB for $L^2(\mathbb{R})$.

Definition. Let H be a separable Hilbert space, $\{e_n\}_{n\in\mathbb{Z}} \subset H$ is a *frame* for H if $\exists A, B > 0$ s.t.

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

If $A = B$, then it is called a *tight frame*.

The Shannon Wavelet Decomposition 3.7 asserts that $\{\psi_{m,n/(4\Omega)}\}$ is a tight wavelet frame with $A = B$.

Proposition 3.8. *If $\{e_n\}$ is a frame for H, then*

$$\forall x \in H, \ x = \sum \langle x, S^{-1}e_n\rangle \cdot e_n, \ \textit{where } Sx = \sum \langle x, e_n\rangle \cdot e_n$$

It turns out that Sx defines a topological isomorphism (I think this is directly from open mapping?). See chapter 12 in [6].

Lecture 4

23 September 2008

Definition. Let $K \subseteq \hat{\mathbb{R}}$ be called τ *congruent* to $[-1/2, 1/2)^d$ iff $\exists\{K_j\}$ a disjoint (except on sets of measure 0) set of Lebesgue measurable subsets of $\hat{\mathbb{R}}^d$, and $\exists\{k_j\} \subseteq \mathbb{Z}^d$ s.t. $\{K_j\}$ is a partition of K and $\{K_j + k_j\}$ is a partition of $[-1/2, 1/2)^d$.

Example 4.1. Let $K = [-1, -1/2) \cup [1/2, 1) \subseteq \hat{\mathbb{R}}$.

1. K is τ congruent to $[-1/2, 1/2)$. Let $K_- = [-1, -1/2)$ and $K_+ = [1/2, 1)$. Then, $(K_- + 1) \cup (K_+ - 1) = [-1/2, 1/2)$. This example is relevant to the artwork of M.C. Escher.

2. $\{2^m K\}_{m \in \mathbb{Z}}$ is a partition of $\hat{\mathbb{R}}$.

3. Let $\hat{\psi} = \mathbb{1}_{[-1,-1/2)} + \mathbb{1}_{[1/2,1)} = \mathbb{1}_K$. Note that if $f \in \mathrm{L}^2(\mathbb{R})$, and $m, n \in \mathbb{Z}$, then

$$\begin{aligned}
\int_{\hat{\mathbb{R}}} \hat{\psi}_{m,n} \cdot \hat{f} d\lambda &= \int_{-1}^{-1/2} \hat{(}2^m\lambda) \cdot e^{2\pi i n\lambda} d\lambda + \int_{1/2}^{1} \hat{f}(2^m\lambda) \cdot e^{2\pi i n\lambda} d\lambda \\
&= \int_{0}^{1/2} \hat{f}(2^m(\gamma - 1)) e^{2\pi i n\gamma} d\gamma + \int_{-1/2}^{0} \hat{f}(2^m(\gamma + 1)) \cdot e^{2\pi i n\gamma} d\gamma \\
&= \int_{-1/2}^{1/2} e^{2\pi i n\gamma} \cdot \Big(\hat{f}(2^m(\gamma - 1))\mathbb{1}_{K_-}(\gamma - 1) + \hat{f}(2^m(\gamma + 1))\mathbb{1}_{K_+}(\gamma + 1)\Big) d\gamma
\end{aligned}$$

This example will reappear in later work, and is instrumental throughout this lecture.

I took the liberty of introducing this as a lemma. It was presented in lecture as part of a theorem.

Lemma 4.2. *If $K \subseteq \hat{\mathbb{R}}^d$ is Lebesgue measurable s.t.*

1. *K is τ congruent to* $[-1/2, 1/2)^d$.

2. $\{2^m K\}$ *is a partition (tiling) of* $\hat{\mathbb{R}}^d$.

3. $|K| = 1$.

Let $\hat{\psi} = \mathbb{1}_K$. Then, $\forall f \in L^2(\mathbb{R}^d)$, we have

$$\sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} |\langle f, \psi_{m,n}\rangle|^2 = \| f \|^2_{L^2(\mathbb{R}^d)}$$

Proof. Obviously, our example works for $d = 1$, but it isn't clear that this even works for higher dimensions (it does). No matter for now, suppose we had such a set. Then, note

$$\begin{aligned}
\sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} |\langle f, \psi_{m,n}\rangle|^2 &= \sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} 2^{md} \left| \int_{\hat{R}^d} \hat{f}(\gamma) \cdot 2^{-md} \cdot e^{2\pi i (n2^{-m})\cdot\gamma} \cdot \overline{\hat{\psi}(2^{-m}\gamma)} d\gamma \right|^2 \text{ by Parseval's theorem} \\
&= \sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} 2^{-md} \left| \int_{\hat{R}^d} \hat{f}(2^m\lambda) \cdot e^{2\pi i n\cdot\lambda} \cdot \overline{\hat{\psi}(\lambda)} 2^{md} d\gamma \right|^2 \\
&= \sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} 2^{md} \left| \int_K \hat{f}(2^m\lambda) \cdot e^{2\pi i n\cdot\lambda} d\lambda \right|^2 \\
&= \sum_{\substack{m\in\mathbb{Z} \\ n\in\mathbb{Z}^d}} 2^{md} \left| \int_{[-1/2,1/2)^d} e^{2\pi i n\cdot\lambda} \underbrace{\left(\sum_j \hat{f}(2^m(\gamma - k_j)) \cdot \mathbb{1}_{K-j}(\gamma - k_j) \right)}_{G_m(\gamma)} d\gamma \right|^2 \\
&= \sum_{m\in\mathbb{Z}} 2^{md} \left(\int_{[-1/2,1/2)^d} \left| G_m(\gamma) \right|^2 d\gamma \right) \text{ by Parseval's theorem} \\
&= \sum_{m\in\mathbb{Z}} 2^{md} \int_K \left| \hat{f}(2^m\lambda) \right|^2 d\lambda \\
&= \sum_{m\in\mathbb{Z}} \int_{2^m K} \left| \hat{f}(\gamma) \right|^2 d\gamma \\
&= \| \hat{f} \|^2_{L^2(\hat{\mathbb{R}}^d)} \text{ by Property 2 of } K \\
&= \| f \|^2_{L^2(\mathbb{R}^d)} \text{ by Plancherel}
\end{aligned}$$

□

Theorem 4.3. *Let $K = [-1, -1/2) \cup [1/2, 1)$ and $\hat{\psi} = \mathbb{1}_K$. Then, $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.*

Proof. The set $\{\psi_{m,n}\}$ is orthonormal (asserted last class, and an immediate consequence of the Plancherel theorem). To conclude that it is a basis, it remains to show that

$$\forall f \in L^2(\mathbb{R}), \ \sum |\langle f, \psi_{m,n}\rangle|^2 = \| f \|^2_{L^2(\mathbb{R})}$$

This follows immediately from the previous lemma. □

Theorem 4.4. *Let $K \subseteq \hat{\mathbb{R}}^d$ be a Lebesgue measurable set s.t.*

1. $|K| = 1$.
2. $\{2^m K\}$ *is a partition (tiling) of* $\hat{\mathbb{R}}^d$.
3. K *is* τ *congruent to* $[-1/2, 1/2)^d$.

Set $\psi = \check{\mathbb{1}}_K$. *Then,* $\{\psi_{m,n}\}$ *forms an ONB for* $L^2(\mathbb{R}^d)$.

Proof. As in the previous, this follows from Plancherel and the lemma. □

Definition. Let $I_{m,n} = \{x \in \mathbb{R} \mid 2^{-m}n \leq x < 2^{-m}(n+1)\}$. This is called a *dyadic interval*. Note that $I_{m+1,2n}$ is the lower half of $I_{m,n}$ and $I_{m+1,2n+1}$ is the upper half.

I have taken liberties with this result, to pick the low hanging fruit for the fact that the Haar wavelets form an orthogonal system.

Lemma 4.5. *Let* $I_{m,n}$ *and* $I_{p,q}$ *be dyadic intervals and assume WLOG that* $m \leq p$. *Then, exactly one of the following holds:*

1. $I_{m,n} = I_{p,q} \iff m = p$ *and* $n = q$.
2. $I_{m,n} \cap I_{p,q} = \emptyset$.
3. $m < p$ *and* $I_{p,q} \subseteq I_{m+1,2n} \subset I_{m,n}$.
4. $m < p$ *and* $I_{p,q} \subseteq I_{m+1,2n+1} \subset I_{m,n}$.

Proof. Suppose not, for a contradiction. The first two cases are trivial, so it suffices to assume that $m < p$. Then, it must be the case that

$$\frac{q}{2^p} < \frac{n}{2^m} = \frac{2n}{2^{m+1}} < \frac{q+1}{2^p} \text{ or}$$
$$\frac{q}{2^p} < \frac{2n+1}{2^{m+1}} < \frac{q+1}{2^p} \text{ or}$$
$$\frac{q}{2^p} < \frac{n+1}{2^m} = \frac{2n+2}{2^{m+1}} < \frac{q+1}{2^p}$$

These all reduce to the case of $M \leq p$, and

$$\frac{q}{2^p} < \frac{N}{2^M} < \frac{q+1}{2^p}$$

But, note that $p - M \geq 0$, so $N \cdot 2^{p-M}$ is an integer, and

$$q < N \cdot 2^{p-M} < q + 1$$

Then, we have found an integer strictly between q and $q + 1$, so we have reached the desired contradiction. Note that the actual conclusion of this is that $I_{p,q}$ is in one half of $I_{m,n}$. □

Theorem 4.6. *Let* $\psi = \mathbb{1}_{[0,1/2} - \mathbb{1}_{[1/2,1)}$ *be the Haar function. Then,*

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

2. *$\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is orthonormal.*

Proof. Part (1.) Note that $x \in \operatorname{supp}\psi_{m,n} \iff 2^m x - n \in [0,1) \iff 2^m x \in [n, n+1) \iff x \in [2^{-m}n, 2^{-m}(n+1)) = I_{m,n}$.

Part (2.) Consider $\langle\psi_{m,n}, \psi_{p,q}\rangle$. From the previous lemma, there are only four cases.
Case (1.) $m = p$ and $n = q$, and (by definition of ψ) we have $\langle\psi_{m,n}, \psi_{p,q}\rangle = \|\psi_{m,n}\|_{L^2(\mathbb{R})} = 1$.
Case (2.) Obviously, $\langle\psi_{m,n}, \psi_{p,q}\rangle = 0$, because $\operatorname{supp}\psi_{m,n} \cap \operatorname{supp}\psi_{p,q} = \emptyset$.
Cases (3.) and (4.) From the lemma, we can conclude that $\psi_{m,n}$ is constant (± 1) on $\operatorname{supp}\psi_{p,q}$. □

Theorem 4.7. *Let ψ be the Haar function. Then, $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.*

Proof. That it is an orthonormal system was just shown. Now, it suffices to show that

$$\forall f \in L^2(\mathbb{R}),\ f = \sum \langle f, \psi_{m,n}\rangle \psi_{m,n} \text{ in } L^2(\mathbb{R}) \text{ norm.}$$

This will be done as follows: Given $\varepsilon > 0$ and $g \in L^2(\mathbb{R})$.
First, show that $\exists N \in \mathbb{N}$, $\exists g_N \in L^2(\mathbb{R})$ s.t. $\operatorname{supp} g_N \subseteq [-2^N, 2^N]$ and

$$\| g - g_N \|_{L^2(\mathbb{R})} < \varepsilon/3.$$

Second, show that $\exists M \in \mathbb{N}$ and $\exists f \in L^2(\mathbb{R})$ s.t. $\operatorname{supp} f \subseteq [-2^N, 2^N]$, f is constant on intervals $I_{m,n}$, and

$$\| g_N - f \|_{L^2(\mathbb{R})} < \varepsilon/3.$$

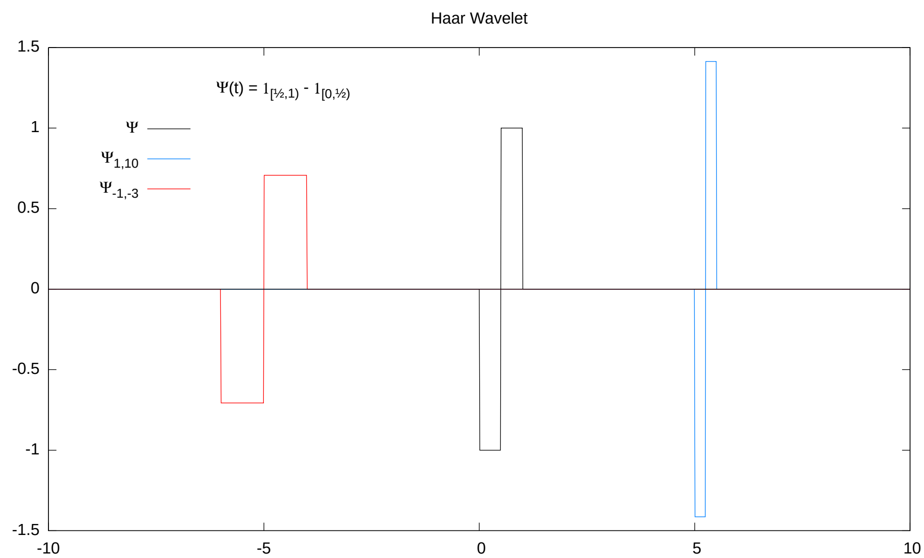
Lastly (and most difficult), $\exists\{c_{m,n} \in \mathbb{C} \mid m,n \in F_1 \times F_2\ F_j \subseteq \mathbb{Z} \text{ finite}\}$ s.t.

$$\| f - \sum c_{m,n}\psi_{m,n} \|_{L^2(\mathbb{R})} < \varepsilon/3.$$

This outline will be completed in two weeks. □

Lecture 5

07 October 2008



Theorem 5.1. *Let ψ be the Haar function. Then, $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.*

Proof. That it is an orthonormal system was shown last lecture. Now, it suffices to show that any function in $\mathrm{L}^2(\mathbb{R})$ can be approximated by a sum of the $\psi_{m,n}$ (in $\mathrm{L}^2(\mathbb{R})$ norm). This will be done as follows: Given $\varepsilon > 0$ and $g \in \mathrm{L}^2(\mathbb{R})$.

1. $\exists N \in \mathbb{N}$, $\exists g_N \in \mathrm{L}^2(\mathbb{R})$ s.t. $\operatorname{supp} g_N \subseteq [-2^N, 2^N]$ and
$$\| g - g_N \|_{\mathrm{L}^2(\mathbb{R})} < \varepsilon/3.$$
2. $\exists M \in \mathbb{N}$ and $\exists f \in \mathrm{L}^2(\mathbb{R})$ s.t. $\operatorname{supp} f \subseteq [-2^N, 2^N]$, f is constant on intervals $I_{m,n}$, and
$$\| g_N - f \|_{\mathrm{L}^2(\mathbb{R})} < \varepsilon/3.$$

3. $\exists\{c_{m,n} \in \mathbb{C} \mid m, n \in F_1 \times F_2 \; F_j \subseteq \mathbb{Z} \text{ finite}\}$ s.t.

$$\| f - \sum c_{m,n}\psi_{m,n} \|_{L^2(\mathbb{R})} < \varepsilon/3.$$

Parts (1.) & (2.) follow from the elementary properties of integration, and will be left as exercises.

Part (3.) We will do this iteratively. For the first step, define $f_0 = f$. Then, we will define f_{-1}, and write $f_0 = f_{-1} + e_{-1}$ (where e_{-1} denotes an error term, and does not conform to our earlier notational conventions).

At this point, Professor Benedetto introduced some notation with brackets that I found confusing rather than enlightening, so I am just not going to omit it. Recall that we defined $f = f_0$ as constant on the intervals $I_{M,n}$. We will define f_{-1} to be constant on intervals $I_{M-1,n}$. On the interval $I_{M-1,n}$, set

$$f_{-1}(t) = \frac{1}{2} \cdot \left(f_0\left(\frac{2n}{2^M}\right) + f_0\left(\frac{2n+1}{2^M}\right)\right),$$

or the average of f_0 on the intervals $I_{M,2n}$ and $I_{M,2n+1}$ (where it was constant on each). So, f_{-1} is constant on intervals twice as long as that of f_0.

By definition, e_{-1} on is $f_0 - f_{-1}$. Then, on the interval $I_{M,2n}$ we see that

$$e_{-1}(t) = f_0\left(\frac{2n}{2^M}\right) - \frac{1}{2} \cdot \left(f_0\left(\frac{2n}{2^M}\right) + f_0\left(\frac{2n+1}{2^M}\right)\right) = \frac{1}{2} \cdot \left(f_0\left(\frac{2n}{2^M}\right) - f_0\left(\frac{2n+1}{2^M}\right)\right)$$

Similarly, on the interval $I_{M,2n+1}$ we see

$$e_{-1}(t) = f_0\left(\frac{2n+1}{2^M}\right) - \frac{1}{2} \cdot \left(f_0\left(\frac{2n}{2^M}\right) + f_0\left(\frac{2n+1}{2^M}\right)\right) = -\frac{1}{2} \cdot \left(f_0\left(\frac{2n}{2^M}\right) - f_0\left(\frac{2n+1}{2^M}\right)\right)$$

Then, it is clear from inspection that on $I_{M-1,n}$ we have

$$e_{-1}(t) = e_{-1}(2k/2^M) \cdot (\mathbb{1}_{I_{M,2n}} - \mathbb{1}_{I_{M,2n+1}}) \implies e_{-1}(t) = 2^{-(M-1)/2} \cdot e_{-1}(2k/2^M) \cdot \psi_{M-1,n} \text{ and}$$

$$e_{-1} = \sum_{k=-2^{N+M-1}}^{2^{N+M-1}-1} 2^{-(M-1)/2} \cdot e_{-1}(2k/2^M) \cdot \psi_{M-1,k} \text{ on } [-2^N, 2^N] \implies f = f_{-1} + \sum_k c_{M-1,k} \cdot \psi_{M-1,k}.$$

Iterate similarly, to define f_{-2} and so forth. This terminates at step $(M + N)$, where $I_{M-(M+N),n} = I_{-N,n}$ of length 2^N. Then, $[-2^N, 2^N] = I_{-N,-1} \cup I_{-N,0}$, and we have

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

where $f_{-(M+N)}$ is constant on the intervals $[-2^n, 0) = I_{-N,-1}$ and $[0, 2^N) = I_{-N,0}$. We reduce notation to call $f_{-(M+N)} = f_{\pm}\mathbb{1}_{[-2^N,2^N)}$. If this function is identically zero, then we are done. That is probably not the case, so we will have to continue this iterative process. We end up approximating f with two functions whose supports are strictly larger support of f. This is unfortunate, but necessary. For any $P > 0$, we get

$$f = 2^{-P} f_{\pm} \mathbb{1}_{[-2^{N+P}, 2^{N+P}} + \underbrace{\sum_{m=1}^{M+N+P} \sum_{k} c_{M-m,k} \psi_{M-m,k}}_{*}$$

Then, we can note that

$$\| f - * \|_{L^2(\mathbb{R})}^2 = \| 2^{-P} f_{\pm} \mathbb{1}_{[-2^{N+P}, 2^{N+P})} \|_{L^2(\mathbb{R})}^2 \leq 2^{-2P} \int_{-2^{N+P}}^{2^{N+P}} \max\{(f\pm)^2\} dt \leq 2^{-2P+N+P+1} \cdot C = 2^{(N+1)-P} \cdot C$$

For sufficiently large P, this will be smaller than $\varepsilon/3$. □

Note that by carefully considering our last few steps, this will work for any $P > 1$, and will not work at all for $P = 1$.

Definition. The *Haar Orthonormal Basis* or *Haar System* for $L^2(\mathbb{T})$ is defined as

$$\{\varphi, \psi_{m,n} \mid m = 0, 1, \ldots, 0 \leq n < 2^m\}, \ \varphi = \mathbb{1}_{[0,1)}.$$

Proposition 5.2. *The Haar System on* $[0, 1)$ *is a basis for* $L^1([0, 1))$.

It is important to note that there does not exist an unconditional basis for $L^1([0, 1))$. This means that you must consider ordered bases, because the associated sums of elements may not converge for arbitrary rearrangements (hence, unconditionally).

At this point, two new homework problems are introduced (which will be incorporated somehow in the homework set):

1. Does the set $\{\tau_k \psi_{m,n} \mid k \in \mathbb{Z}, m = 0, 1 \ldots, 0 \leq n < 2^m\}$ form a basis for $L^1(\mathbb{R})$?
2. Let $f \in L^p(\mathbb{T})$, with $1 < p < \infty$, with wavelet coefficients $\langle f, \psi_{m,n} \rangle$. Suppose that $g \in L^1(\mathbb{T})$ has the property that $\forall m = 0, 1 \ldots, 0 \leq n < 2^m$ we have $|\langle f, \psi_{m,n} \rangle| = |\langle g, \psi_{m,n} \rangle|$. Prove that $g \in L^p(\mathbb{T})$. What can you say about the relation between g and f?

Proposition 5.3 (Some Haar System Properties)**.** *Choose an ordering for the Haar system for* $\mathbb{T}, \theta_p$ *for* $p = 0, 1, \ldots$. *Define the Haar series of* $f \in L^1(\mathbb{T})$

$$H(f)(\gamma) = \sum_{p=0}^{\infty} f[p] \theta_p(\gamma), \ \textit{where} \ f[p] = \int_{\mathbb{T}} f(t) \theta_p(t) dt$$

Then, in contrast to Fourier Series, for $F \in L^1(\mathbb{T})$

- $H(F) = F$ *a.e.*
- F *continuous at* $\gamma \implies H(F)(\gamma) = F(\gamma)$.

There are several other similar type results regarding F and $H(F)$ which were mentioned, but not spelled completely out. While these results indicate significant improvements for Haar series versus Fourier series, this is not always the case. Also in contrast to Fourier series,

- *If $F \in C(\mathbb{T})$ and $\sum |f[p]| < \infty$, then it is* not *the case that $\sum f[p]\theta_p$ converges uniformly.*

Now, returning to the definition of an MRA.

Definition. The pair $\{V_j\}_{j\in\mathbb{Z}}, \varphi$ is an *Multi Resolution Analysis* of $\mathrm{L}^2(\mathbb{R})$ if

1. each V_j is a closed (vector) subspace (of $\mathrm{L}^2(\mathbb{R})$.
2. $V_j \subseteq V_{j+1}$ for each $j \in \mathbb{Z}$.
3. $\bigcup V_j = \mathrm{L}^2(\mathbb{R})$ and $\bigcap V_j = 0$ (the zero function).
4. $f(t) \in V_j \iff f(2t) \in V_{j+1}$.
5. $f \in V_0 \iff \forall k \in \mathbb{Z}, \tau_k f \in V_0$.
6. $\varphi \in V_0$ and $\{\tau_k\varphi\}_{k\in\mathbb{Z}}$ for an orthonormal basis (ONB) for V_0.

The φ is called a *scaling function*.

Proposition 5.4. *Let $\{V_j\}, \varphi$ be an MRA of $L^2(\mathbb{R})$. Then, there is a constructible function ψ, which depends on φ, s.t. $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$.*

Examples of this are the Haar (to be shown) and Shannon systems.

Let $\varphi \in \mathrm{L}^2(\mathbb{R})$, and define

$$\Phi(\gamma) = \sum_{n\in\mathbb{Z}} |\hat{\varphi}(\gamma + n)|^2$$

Note that $\Phi \in \mathrm{L}^1(\mathbb{T})$ by the Beppo-Levi and Plancherel theorems.

Theorem 5.5. *Let $\varphi \in L^2(\mathbb{R})$ and let $V = \overline{span}\{\tau_k\varphi\}$. Then, $\{\tau_n\varphi\}$ is an ONB for V if and only if $\Phi = 1$ a.e.*

Proof. Consider

$$\begin{aligned}\int \varphi(t)\overline{\varphi(t-n)}dt &= \int |\hat{\varphi}(\gamma)|^2 e^{2\pi i n\gamma} d\gamma \\ &= \sum_k \int_{\mathbb{T}} |\hat{\varphi}(\gamma + k)|^2 \cdot e^{2\pi i n\gamma} d\gamma \\ &= \int_T \Phi(\gamma) \cdot e^{2\pi i n\gamma} d\gamma\end{aligned}$$

There are several standard analysis results that justifies interchanging the sum and the integral in the last step of the above, namely the dominated convergence theorem or the Weierstraß M-test.

From the basic facts from Fourier series, the left hand side of the above is $\delta_{n,0} \iff \Phi$ is 1 a.e. □

There are some other results similar in nature to the above:

- $\{\tau_n \varphi\}$ is a Riesz basis for $V \iff \exists A, B > 0$ s.t. $A \leq \Phi \leq B$ a.e. on $\mathbb{T}$
- $\{\tau_n \varphi\}$ is a frame for $V \iff \exists A, B > 0$ s.t. $A \leq \Phi \leq B$ almost everywhere that Φ is non-zero.

Definition. Suppose that $H \in \mathrm{L}^1(\mathbb{T})$. H is called a *conjugate mirror filter* or *quadrature mirror filter* if $|H(\gamma)|^2 + |H(\gamma + 1/2)|^2 = 2$ a.e. Professor Benedetto cited a book by author P.P. Vaidyanathan. I think [15] must be the one he intended.

Proposition 5.6. *Let $\varphi \in L^2(\mathbb{R})$ and assume that $\{\tau_n \varphi\}$ is ON and*

$$\hat{\varphi}(\gamma) = \frac{1}{\sqrt{2}} \cdot H_0\left(\frac{\gamma}{2}\right)\hat{\varphi}\left(\frac{\gamma}{2}\right) \text{ a.e. for } H_0 \in L^1(\mathbb{T})$$

Then, H_0 is a conjugate mirror filter.

Lecture 6

14 October 2008

Recall the definition $H \in \mathrm{L}^1(\mathbb{T})$ is a *conjugate mirror filter* if $|H(\gamma)|^2 + |H(\gamma+1/2)|^2 = 2$ a.e. Note that such an $H \in \mathrm{L}^\infty(\mathbb{T})$ and $\| H \|_{\mathrm{L}^\infty(\mathbb{T})} \leq \sqrt{2}$. Also, define the $A(\mathbb{T}) \subseteq C(\mathbb{T})$ to the space of absolutely convergent Fourier series.

Proposition 6.1. *Let $H \in A(\mathbb{T})$ and*

$$h[n] = \int_{\mathbb{T}} H(\gamma) \cdot e^{2\pi i n \gamma} d\gamma,$$

then H is a conjugate mirror filter if and only if

$$\sum_{n \in \mathbb{Z}} h[n-2j] \cdot \overline{h[n-2k]} = \delta(j,k) = \begin{cases} 1 & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

In this case, $\| H \|_{L^2(\mathbb{T})} = 1$ and

$$\| H \|^2_{L^2(\mathbb{T})} \leq \| H \|^2_{L^\infty(\mathbb{T})} \leq \sum_{n \in \mathbb{Z}} |h[n]| = \|H\|_A$$

Definition. Let $U = (u_{jk})$ be an element of $N \times N$ complex matrices. U is *unitary* if and only if its column (or row) vectors are orthonormal.

Proposition 6.2. *Let $H_0, H_1 \in L^1(\mathbb{T})$. Define*

$$U(\gamma) = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} H_0(\gamma) & H_0(\gamma + 1/2) \\ H_1(\gamma) & H_1(\gamma + 1/2) \end{pmatrix} \tag{6.1}$$

U is unitary a.e. $\iff$ H_0, H_1 are conjugate mirror filters and

$$H_0(\gamma) \cdot \overline{H_1(\gamma)} + H_0(\gamma + 1/2) \cdot \overline{H_1(\gamma + 1/2)} = 0 \text{ a.e.}$$

Theorem 6.3. *Let $H_0, H_1 \in L^1(\mathbb{T})$, define U as in eq. 6.1. Then, U is unitary if and only if H_0, H_1 are CMFs and $\exists K$ 1-periodic s.t.*

$$H_1(\gamma) = K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}, \; K(\gamma + 1/2) + K(\gamma) = 0, \; |K| = 1$$

Proof. "$\Longrightarrow$". Suppose U is unitary. Then,

$$H_1(\gamma) \cdot \overline{H_0(\gamma)} = -H_1(\gamma + 1/2) \cdot \overline{H_0(\gamma + 1/2)} \tag{6.2}$$

Let $A = \{\gamma \mid H_0(\gamma) = 0\}$, $B = \{\gamma \mid H_0(\gamma) \neq 0\}$. Then, $A \cap B = \emptyset$ and $A \cup B = \mathbb{T}$. If $\gamma \in B$, then define

$$K(\gamma) = \frac{-H_1(\gamma + 1/2)}{\overline{H_0(\gamma)}}$$

If $|A| > 0$ (or it doesn't matter), and $\gamma \in A$, define

$$K(\gamma) = \frac{H_1(\gamma)}{\overline{H_0(\gamma + 1/2)}}$$

Then, using eq. 6.2, we see that $\forall \gamma \in \mathbb{T},\ H_1(\gamma) = K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}$.
To see that $K(\gamma+1/2) = -K(\gamma)$ entails three similar cases. To demonstrate one, suppose that $\gamma, \gamma + 1/2 \in B$, then note that

$$K(\gamma + 1/2) = \frac{-H_1((\gamma + 1/2) + 1/2)}{\overline{H_0(\gamma + 1/2)}} = \frac{-H_1(\gamma)}{\overline{H_0(\gamma + 1/2)}} = \frac{H_1(\gamma + 1/2)}{\overline{H_0(\gamma)}} = -K(\gamma)$$

The case where $\gamma \in A,\ \gamma + 1/2 \in B$ or $\gamma \in B,\ \gamma + 1/2 \in A$ follow along similar lines (one less step).

Finally, to see that $|K| = 1$ (left out in class, because it's simple and tedious). Note that

$$H_1(\gamma) = K(\gamma) \cdot \overline{H_0(\gamma + 1/2)} \text{ and } K(\gamma + 1/2) = -K(\gamma) \implies$$
$$H_1(\gamma + 1/2) = K(\gamma + 1/2) \cdot \overline{H_0(\gamma)} = -K(\gamma)\overline{H_0(\gamma)}$$

By assumption, U is unitary, and note that

$$1 = \langle U_2(\gamma), U_2(\gamma) \rangle = \frac{1}{2} \cdot \Big\langle (H_1(\gamma), H_1(\gamma + 1/2)), (H_1(\gamma), H_1(\gamma + 1/2)) \Big\rangle =$$
$$\frac{1}{2} \cdot \Big\langle (K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}, -K(\gamma)\overline{H_0(\gamma)}), (K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}, -K(\gamma)\overline{H_0(\gamma)}) \Big\rangle =$$
$$\frac{1}{2} \cdot \Big(|H_0(\gamma)|^2 + |H_0(\gamma + 1/2)|^2 \Big) \cdot |K(\gamma)|^2 = |K(\gamma)|^2$$

"$\Longleftarrow$". The row vectors of $U(\gamma)$ have norm 1, by the CMF hypothesis. Note that using $H_1(\gamma) = K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}$, then,

$$U(\gamma) = \frac{1}{\sqrt{2}} \cdot \begin{pmatrix} H_0(\gamma) & H_0(\gamma + 1/2) \\ K(\gamma) \cdot \overline{H_0(\gamma + 1/2)} & K(\gamma + 1/2) \cdot \overline{H_0(\gamma)} \end{pmatrix}$$

Note that

$$\begin{aligned}
\langle U_1(\gamma), U_2(\gamma) \rangle &= \Big\langle (H_0(\gamma), H_0(\gamma + 1/2)), (K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}, K(\gamma + 1/2) \cdot \overline{H_0(\gamma)}) \Big\rangle \\
&= \overline{K(\gamma)} \cdot H_0(\gamma) \cdot H_0(\gamma + 1/2) + \overline{K(\gamma + 1/2)} \cdot H_0(\gamma) \cdot H_0(\gamma + 1/2) \\
&= H_0(\gamma) \cdot H_0(\gamma + 1/2) \cdot (\underbrace{\overline{K(\gamma) + K(\gamma + 1/2)}}_{0}) \\
&= 0
\end{aligned}$$

Then, U is unitary, since its row vectors are orthonormal. $\square$

Remark: Suppose that $H_0, H_1 \in L^\infty(\mathbb{T})$ with H_0 a CMF, $H_1(\gamma) = K(\gamma) \cdot \overline{H_0(\gamma + 1/2)}$ and $K(\gamma) + K(\gamma + 1/2) = 0$. Embedded in the previous proof is the fact that

$$|H_1(\gamma)|^2 + |H_1(\gamma + 1/2)|^2 = 2|K(\gamma)|^2.$$

Then, the fact that $|K(\gamma)| = 1$ is critical to the conclusion. For example, consider $K(\gamma) = \sin(2\pi\gamma)$.

Corollary 6.4. *Let H_0 be a CMF and $F \in L^1(\mathbb{T})$. Then,*

$$F(\gamma) \cdot \overline{H_0(\gamma)} = -F(\gamma + 1/2) \cdot \overline{H_0(\gamma + 1/2)} \iff F(\gamma) = K_F(\gamma) \cdot \overline{H_0(\gamma + 1/2)}$$

for some 1-periodic K_F satisfying $K_F(\gamma) + K_F(\gamma + 1/2) = 0$.

Example 6.5. Given H_0, H_1 as in the theorem, assume that their Fourier coefficients are related by

$$h_1[n] = (-1)^n \overline{h_0[-n+1]}$$

Then, $K(\gamma) = -e^{-2\pi i\gamma}$. In fact,

$$\begin{aligned} H_1(\gamma + 1/2) &= \sum_{k\in\mathbb{Z}} h_1[k] e^{-2\pi i k(\gamma+1/2)} \\ &= \sum_{k\in\mathbb{Z}} \overline{h_0[-k+1]} e^{-2\pi i k\gamma} \\ &= e^{-2\pi i\gamma} \overline{H_0(\gamma)} \end{aligned}$$

This equality holds a.e, by virtue of Lusin-Carleson.

Lemma 6.6. *Let $\varphi \in L^2(\mathbb{R})$. Assume that $\{\tau_n\varphi\}$ is an orthonormal system and*

$$\hat{\varphi}(\gamma) = \frac{1}{\sqrt{2}} \cdot H_0\left(\frac{\gamma}{2}\right) \cdot \hat{\varphi}\left(\frac{\gamma}{2}\right)$$

for some $H_0 \in L^1(\mathbb{T})$. Then, H_0 is a CMF (and hence $H_0 \in L^\infty(\mathbb{R})$).

Proof. Recall that we proved last lecture (theorem 5.5) that

$$\{\tau_n\varphi\} \text{ is ON} \iff \sum_{n\in\mathbb{Z}} |\hat{\varphi}(\gamma + n)|^2 = \Phi(\gamma) = 1 \text{ a.e.}$$

Note that

$$\hat{\varphi}(\gamma+n) = \frac{1}{\sqrt{2}} \cdot H_0\Big(\frac{\gamma}{2}+\frac{n}{2}\Big)\cdot \hat{\varphi}\Big(\frac{\gamma}{2}+\frac{n}{2}\Big) \implies$$

$$1 = \Phi(\gamma) = \frac{1}{2}\sum_{n\in\mathbb{Z}} \left|H_0\Big(\frac{\gamma}{2}+\frac{n}{2}\Big)\right|^2 \cdot \left|\hat{\varphi}\Big(\frac{\gamma}{2}+\frac{n}{2}\Big)\right|^2 \text{ letting } \lambda = \frac{\gamma}{2} \implies$$

$$2 = \sum_{n\in\mathbb{Z}} \left|H_0\Big(\lambda+\frac{n}{2}\Big)\right|^2 \cdot \left|\hat{\varphi}\Big(\lambda+\frac{n}{2}\Big)\right|^2$$

$$= \sum_{\substack{n\in\mathbb{Z}\\ n\text{ even}}} \left|H_0\Big(\lambda+\frac{n}{2}\Big)\right|^2 \cdot \left|\hat{\varphi}\Big(\lambda+\frac{n}{2}\Big)\right|^2 + \sum_{\substack{n\in\mathbb{Z}\\ n\text{ odd}}} \left|H_0\Big(\lambda+\frac{n}{2}\Big)\right|^2 \cdot \left|\hat{\varphi}\Big(\lambda+\frac{n}{2}\Big)\right|^2$$

$$= |H_0(\lambda)|^2 \underbrace{\sum_{m\in\mathbb{Z}} |\hat{\varphi}(\lambda+m)|^2}_{\Phi(\gamma)=1} + |H_0(\lambda+1/2)|^2 \underbrace{\sum_{m\in\mathbb{Z}} |\hat{\varphi}(\lambda+1/2+m)|^2}_{\Phi(\gamma+1/2)=1}$$

□

Assume that $\{V_j\}$ is an MRA of $\mathbb{R}$ with scaling function φ. Define W_j to be the orthogonal complement of V_j in V_{j+1}. Then, W_j is a closed subspace of V_{j+1} and $V_{j+1} = V_j \oplus W_j$ (i.e. $f \in V_{j+1}$ decomposes uniquely as $f = f_V + f_W$ with $f_V \in V_J$, $f_W \in W_j$ and $f_V \perp f_W$).

Remark: Note that $f(t) \in W_j$ if and only if $f(t/2) \in W_{j-1}$. To see this, suppose that $f \in W_m$. Then, $f \in V_{m+1} \supseteq W_m$. Therefore, $f(t/2) \in V_m = V_{m-1} \oplus W_{m-1}$. On the other hand, $f(t/2) \perp V_{m-1}$ because

$$\langle \varphi_{m,n}, f(t/2)\rangle = 2^{(m-1)/2} \int \varphi(2^{m-1}t-n)\cdot \frac{1}{\sqrt{2}} \cdot \overline{f\Big(\frac{t}{2}\Big)} dt \text{ (letting } u = t/2)$$

$$= 2^{m/2} \int \underbrace{\varphi(2^m u - n)}_{\in V_m} \cdot \underbrace{\overline{f(u)}}_{\in W_m} du = 0$$

Lemma 6.7. *Given the MRA $\{V_j\}$ with scaling function φ, there is a canonically constructed CMF associated with this MRA.*

Proof. Consider that $\varphi_{-1,0} = (1/\sqrt{2})\varphi(t/2) \in V_{-1} \subset V_0$. The set $\{\tau_n\varphi\}$ is an ONB for V_0. Then, we know that

$$\frac{1}{\sqrt{2}} \cdot \varphi\Big(\frac{t}{2}\Big) = \sum_{n\in\mathbb{Z}} h_0[n]\cdot \varphi(t-n) \tag{6.3}$$

where

$$\{h_0[n]\} \in \ell^2(\mathbb{Z}) \text{ and } h_0[n] = \int_{\mathbb{R}} \varphi_{-1,0}(t)\cdot \overline{\varphi(t-n)}dt.$$

Taking the Fourier Transform of eq. 6.3, we get

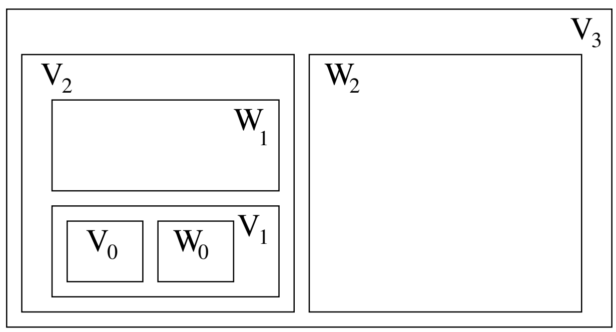
$$\sqrt{2}\cdot\hat{\varphi}(2\gamma) = H_0(\gamma)\cdot\hat{\varphi}(\gamma), \text{ where } H_0(\gamma) = \sum_{n\in\mathbb{Z}} h_0[n]\cdot e^{-2\pi i n\gamma} \in L^2(\mathbb{T}) \subseteq L^1(\mathbb{T})$$

From the previous lemma, we can conclude that H_0 is a CMF. This is apparently referred to as the *solution of the frequency scaling equation*, for the obvious reason. $\square$

In the class, Professor Benedetto began the proof of the MRA theorem. I moved proof to the next class for completeness.

Lecture 7

21 October 2008



Theorem 7.1 (MRA Theorem)**.** *Let $\{V_j\}$ be an MRA of $L^2(\mathbb{R})$ with scaling function $\varphi \in L^2(\mathbb{R})$. There is a canonically constructible $\psi \in W_0$ for this MRA s.t. $\{\psi_{m,n}\}$ is an ONB for $L^2(\mathbb{R})$. In fact, if H_0 is the solution to the frequency scaling equation with Fourier coefficients $h_0[n]$, then*

$$\psi(t) = \sqrt{2}\sum_{n\in\mathbb{Z}} h_1[n]\cdot\varphi(2t-n) \in L^2(\mathbb{R})$$
$$h_1[n] = (-1)^{n+1}\cdot\overline{h_0[-n-1]}$$

Proof. We will break this proof into sections.
Part (a.) Recall that V_1 is the closure of the subspace of $\mathrm{L}^2(\mathbb{R})$ spanned by the collection $\{\tau_n\varphi(2t)\}$. Then,

$$f \in V_1 \iff f = \sum_{n\in\mathbb{Z}}\langle f,\ \sqrt{2}\cdot\varphi(2t-n)\rangle\cdot\ \sqrt{2}\cdot\varphi(2t-n) \iff$$
$$\hat{f} = \sum_{n\in\mathbb{Z}}\langle f,\ \sqrt{2}\cdot\varphi(2t-n)\rangle\cdot\frac{1}{\sqrt{2}}\cdot\hat{\varphi}\left(\frac{\gamma}{2}\right)\cdot e^{-2\pi i(n/2)\gamma} = \frac{1}{\sqrt{2}}\cdot\hat{\varphi}\left(\frac{\gamma}{2}\right)\cdot F_f\left(\frac{\gamma}{2}\right)$$

where

$$F_f\left(\frac{\gamma}{2}\right) \stackrel{\text{def}}{=} \sum_{n\in\mathbb{Z}}\langle f,\ \sqrt{2}\cdot\varphi(2t-n)\rangle\cdot e^{-2\pi i(n/2)\gamma}$$

and F_f is 1-periodic. At this point, we want to establish conditions on φ to ensure that F_f is in a "good" space.
Aside: Note that

$$\langle f, 2^{1/2}\cdot\varphi(2t-n)\rangle = \langle \hat{f}, 2^{-1/2}\cdot e^{-2\pi i(n/2)\gamma}\cdot\hat{\varphi}(\gamma/2)\rangle = \frac{1}{\sqrt{2}}\cdot\int \hat{f}(\gamma)\cdot\hat{\varphi}(\gamma/2)\cdot e^{-2\pi i(n/2)\gamma}d\gamma$$

If $\hat{f},\hat{\varphi}\in L^2(\hat{\mathbb{R}})$, then $\hat{f}(\gamma)\cdot\hat{\varphi}(\gamma/2)\in L^1(\mathbb{R})$ and

$$\frac{1}{\sqrt{2}}\int \hat{f}(\gamma)\cdot\hat{\varphi}(\gamma/2)\cdot e^{-2\pi i t\gamma}d\gamma \in A(\mathbb{R})\subset C_0(\mathbb{R})$$

which is a step towards establishing the control we need, but more work needs to be done (not covered in class).

Part (b.) Assuming that $f\in V_1$, recall that $f\in W_0 \iff f\perp V_0$. Let's show

$$f\in W_0 \iff \sum_{k\in\mathbb{Z}}\hat{f}(\gamma+k)\cdot\overline{\hat{\varphi}(\gamma+k)} = 0 \text{ a.e.} \tag{7.1}$$

Let's begin by noting

$$\begin{aligned} f\perp V_0 &\iff \forall n,\ \int f(t)\cdot\overline{\varphi(t-n)}dt = 0\\ &\iff \forall n,\ \int \hat{f}(\gamma)\cdot\overline{\hat{\varphi}(\gamma)}\cdot e^{2\pi i n\gamma}d\gamma = 0\\ &\iff \forall n,\ \sum_{k\in\mathbb{Z}}\Big[\int_0^1 \hat{f}(\gamma+k)\cdot\overline{\hat{\varphi}(\gamma+k)}\cdot e^{2\pi i n\gamma}d\gamma\Big] = 0\\ &\iff \forall n,\ \int_0^1 \underbrace{\Big[\sum_{k\in\mathbb{Z}}\hat{f}(\gamma+k)\cdot\overline{\hat{\varphi}(\gamma+k)}\Big]}_{G} e^{2\pi i n\gamma}d\gamma = 0 \end{aligned} \tag{7.2}$$

The function G is 1-periodic and

$$\begin{aligned} \|G\|_{L^1(\mathbb{T})} &\le \int_0^1\Big(\sum_{k\in\mathbb{Z}}\big|\hat{f}(\gamma+k)\cdot\hat{\varphi}(\gamma+k)\big|\Big)\\ &= \int_{\hat{\mathbb{R}}}\hat{f}(\gamma)\cdot\hat{\varphi}(\gamma)d\gamma\\ &\le \|\hat{f}\|_{L^2(\hat{\mathbb{R}})}\cdot\|\hat{\varphi}\|_{L^2(\hat{\mathbb{R}})}, \end{aligned}$$

which is finite. Then, $G\in L^1(\mathbb{T})$, so eq. 7.2 $\implies$ $G=0$ a.e. Then, by the uniqueness theorem for Fourier series, we see 7.1 holds, as desired.

Part (c.) Now, let's show that (assuming $f\in V_1$)

$$f\in W_0 \iff F_f\cdot\overline{H_0} + (\tau_{-1/2}F_f)\cdot\overline{(\tau_{-1/2}H_0)} = 0 \tag{7.3}$$

Let's begin by substituting $\hat{\varphi}(\gamma) = 2^{-1/2} \cdot \hat{\varphi}(\gamma/2) \cdot H_0(\gamma/2)$ into 7.1. We see that

$$\begin{aligned}
0 &= \sum_{k\in\mathbb{Z}} \hat{f}(\gamma + k) \cdot \overline{\hat{\varphi}(\gamma + k)} \\
&= \frac{1}{2} \sum_{k\in\mathbb{Z}} F_f\Big(\frac{\gamma}{2} + \frac{k}{2}\Big) \cdot \overline{H_0\Big(\frac{\gamma}{2} + \frac{k}{2}\Big)} \cdot \Big|\hat{\varphi}\Big(\frac{\gamma}{2} + \frac{k}{2}\Big)\Big|^2 \\
&= \frac{1}{2}\Big[\sum_{\substack{k\in\mathbb{Z} \\ k \text{ even}}} (*) + \sum_{\substack{k\in\mathbb{Z} \\ k \text{ odd}}} (*)\Big] \\
&= \Big(F_f(\lambda) \cdot \overline{H_0(\lambda)}\Big)\underbrace{\Big[\sum_{k\in\mathbb{Z}} |\hat{\varphi}(\lambda + k)|^2 \Big]}_{\Phi(\lambda)=1} + \Big(F_f\Big(\lambda + \frac{1}{2}\Big) \cdot \overline{H_0\Big(\lambda + \frac{1}{2}\Big)}\Big)\underbrace{\Big[\sum_{k\in\mathbb{Z}} \Big|\hat{\varphi}\Big(\lambda + k + \frac{1}{2}\Big)\Big|^2 \Big]}_{\Phi(\lambda+1/2)=1}
\end{aligned}$$

Then, we are done, by the previous two lemmas.

Part (d.) Assume that $f \in V_1$. Recall from part (a) that

$$\hat{f}(\gamma) = \frac{1}{\sqrt{2}} \cdot \hat{\varphi}\Big(\frac{\gamma}{2}\Big) \cdot F_f\Big(\frac{\gamma}{2}\Big)$$

Now, using eq. 7.3 and Corollary 6.4, we see that

$$f \in W_0 \iff \hat{f}(\gamma) = \frac{1}{\sqrt{2}} \cdot \hat{\varphi}\Big(\frac{\gamma}{2}\Big) \cdot K_f\Big(\frac{\gamma}{2}\Big) \cdot \overline{H_0\Big(\frac{\gamma}{2} + \frac{1}{2}\Big)}$$

for a 1-periodic K_f with $K_f(\gamma) + K_f(\gamma + 1/2) = 0$.

Note that for any 1-periodic function K, we can define $L(\gamma) = e^{-\pi i\gamma} \cdot K(\gamma/2)$, and L is one-periodic iff $K(\gamma) + K(\gamma + 1/2) = 0$.

Combining this with the above, we have $f \in W_0$ if and only if

$$\hat{f}(\gamma) = \frac{1}{\sqrt{2}} \cdot \hat{\varphi}\Big(\frac{\gamma}{2}\Big) \cdot e^{\pi i\gamma} \cdot L_f(\gamma) \cdot \overline{H_0\Big(\frac{\gamma}{2} + \frac{1}{2}\Big)}$$

where L_f is 1-periodic.

Part (e.) Now, we're almost done, using the above, we can see what ψ should be. Define

$$\hat{\psi}(\gamma) = \frac{1}{\sqrt{2}} \cdot \hat{\varphi}\Big(\frac{\gamma}{2}\Big) \cdot e^{\pi i\gamma} \cdot \overline{H_0\Big(\frac{\gamma}{2} + \frac{1}{2}\Big)} \in \mathrm{L}^2(\hat{\mathbb{R}}). \tag{7.4}$$

Obviously, $\hat{f}(\gamma) = \hat{\psi}(\gamma)L_f(\gamma)$, and L_f can be expanded in a Fourier series (since it's 1-periodic):

$$L_f(\gamma) = \sum_{n\in\mathbb{Z}} \hat{c}_{f,n} e^{2\pi in\gamma}$$

It follows that if $f \in W_0$, then

$$\hat{f}(\gamma) = \sum_{n\in\mathbb{Z}} \hat{c}_{f,n} e^{2\pi i n\gamma} \hat{\psi}(\gamma)$$

$$f = \sum_{n\in\mathbb{Z}} c_{f,n} \cdot (\tau_n \psi)$$

If we show that $\{\tau_n\psi\}$ is ON, then it forms an ONB for W_0. To this, we calculate

$$\int_{\mathbb{R}} \psi(t-m) \cdot \overline{\psi(t-n)} dt = \int_0^1 \Psi(\gamma) \cdot e^{2\pi i(n-m)\gamma} d\gamma$$
$$\Psi(\gamma) = \sum_{n\in\mathbb{Z}} |\hat{\psi}(\gamma+k)|^2$$

We will be done if we show that $\Psi(\gamma) = 1$ a.e. on $[0, 1)$. Expand as

$$\begin{aligned}
\Psi(\gamma) &= \sum_{k\in\mathbb{Z}} |\hat{\psi}(\gamma+k)|^2 \\
&= \sum_{k\in\mathbb{Z}} \left| \frac{1}{\sqrt{2}} \cdot \hat{\varphi}\left(\frac{\gamma}{2} + \frac{k}{2}\right) \cdot e^{\pi i(\gamma+k)} \cdot \overline{H_0\left(\frac{\gamma}{2} + \frac{k}{2} + \frac{1}{2}\right)} \right|^2 \\
&= \frac{1}{2} \cdot \sum_{k\in\mathbb{Z}} \left| \hat{\varphi}\left(\frac{\gamma}{2} + \frac{k}{2}\right) \right|^2 \cdot \left| H_0\left(\frac{\gamma}{2} + \frac{k}{2} + \frac{1}{2}\right) \right|^2
\end{aligned}$$

Using our trick of expanding the even and odd cases, note that we get

$$\begin{aligned}
\Psi(\gamma) &= \frac{1}{2} \cdot \left(\Phi(\gamma) \cdot |H_0(\gamma)|^2 + \Phi\left(\gamma + \frac{1}{2}\right) \cdot \left| H_0(\gamma + \frac{1}{2}) \right|^2 \right) \\
&\varphi \text{ is an ONB, so } \Phi = 1 \; a.e. \\
&= \frac{1}{2} \cdot \left(|H_0(\gamma)|^2 + \left| H_0\left(\gamma + \frac{1}{2}\right) \right|^2 \right) \\
&= 1 \qquad \text{since } H_0 \text{ is a CMF}
\end{aligned}$$

So, we know that $\{\tau_n\psi\}$ forms an ONB for W_0.

To refine the definition of ψ slightly, recall Example 6.5 and define H_1

$$\begin{aligned}
H_1(\gamma) &= e^{2\pi i\gamma} \cdot \overline{H_0\left(\gamma + \frac{1}{2}\right)} \implies \\
h_1[n] &= (-1)^{n+1} \cdot \overline{h_0[-n-1]}, \\
\sqrt{2} \cdot \hat{\psi}(2\gamma) &= \hat{\varphi}(\gamma) \cdot H_1(\gamma), \text{ and} \\
\psi(t) &= \sum_{n\in\mathbb{Z}} h_1[n] \cdot \sqrt{2} \cdot \varphi(2t-n)
\end{aligned}$$

Now, we need to get from W_0 to all of $L^2(\mathbb{R})$.

Part (f.) Recall that $f \in W_0 \iff f(2t) \in W_1$. This extends in the obvious way as a natural isometry $D_m \colon W_0 \to W_m$. Note that $D_m(\psi_{0,n}) = \psi_{m,n}$, so the set $\{\psi_{m,n}\}$ is an ONB for W_m. So, we're making progress.

Part (g.) Definition. Let $\{U_m; m \in \mathbb{Z}\}$ be a sequence of closed linear subspaces of a Banach space B and let

$$X = \{f = \sum_{m \in F(f)} f_m, \quad f_m \in U_m\} \qquad F(f) \text{ is a finite set} \tag{7.5}$$

X is the **direct sum** of $\{U_m\}$, written $\bigoplus U_m$ if each element of X has a unique representation in terms of 7.5.

Note that $V_m = V_{m-1} \oplus W_{m-1}$, and $W_j \subseteq V_{m-1}$ for $j < m$. Then $W_j \perp W_m$ for all $j \neq m$. This implies immediately that if $M \subset \mathbb{N}$ is a finite set, then

$$\operatorname{span}\{W_m\}_{m \in M} = \bigoplus_{m \in M} W_m$$

since if $\sum f_m = \sum g_m$ where $f_m, g_m \in W_m$ and all but finitely many are zero, then

$$\forall m, \ \sum_{j \neq m} (g_j - f_j) = f_m - g_m \in W_m$$

$$\Rightarrow \|f_m - g_m\|^2 = \sum_{j \neq m} \langle g_j - f_j, f_m - g_m \rangle = 0$$

$$\Rightarrow f_m = g_m$$

Part (h.) Extending this, we see that $\forall J \in \mathbb{Z}$

$$\mathcal{W}_J = \overline{\bigoplus_{j=-\infty}^{J} W_j} = \{f \in L^2(\mathbb{R}) \mid \forall \epsilon > 0, \exists g_\epsilon \in \bigoplus_{j=-\infty}^{J} W_j \ s.t. \ \| f - g_\epsilon \|_{L^2(\mathbb{R})} < \epsilon\}.$$

We will show that $\mathcal{W}_J = V_{J+1}$. By our construction, we have that $W_j \subseteq V_{J+1}$ for all $j \leq J$. Then $\bigoplus_{j=-\infty}^{J} W_j \subseteq V_{J+1}$. Since V_{J+1} is closed (by an MRA assumption), $\mathcal{W}_J \subseteq V_{J+1}$.

Pick $f \in V_{J+1}$. Then, $\exists f_J \in V_J, g_J \in W_J$ unique, so that $f = f_J + g_J$. Because $f_J \perp g_J$, we know $\| f \|_{L^2(\mathbb{R})}^2 = \| f_J \|_{L^2(\mathbb{R})}^2 + \| g_J \|_{L^2(\mathbb{R})}^2$, since this is basically equation (4) in [14]'s Theorem 4.11. Continue this with unique $f_{J-1} \in V_{J-1}, g_{J-1} \in W_{J-1}$ so that $f_J = f_{J-1} + g_{J-1}$. Then, $f = f_{J-1} + g_J + g_{J-1}$ and

$$\| f \|_{L^2(\mathbb{R})}^2 = \| f_{J-1} \|_{L^2(\mathbb{R})}^2 + \| g_J \|_{L^2(\mathbb{R})} + \| g_{J-1} \|_{L^2(\mathbb{R})}^2$$

We can continuing recursively. We get $\{f_m\}, \{g_m\}$ s.t. $f_m \in V_m$, $g_m \in W_m$, and $\forall m \in \mathbb{Z}$ with $m \leq J$

$$f = f_m + \sum_{j=m}^{J} g_j$$

$$\| f \|^2_{\mathrm{L}^2(\mathbb{R})} = \| f_m \|^2_{\mathrm{L}^2(\mathbb{R})} + \sum_{j=m}^{J} \| g_j \|^2_{\mathrm{L}^2(\mathbb{R})} \tag{7.6}$$

Consider then the element

$$g = \sum_{j=-\infty}^{J} g_j,$$

which can be written

$$g = \sum_{j=-\infty}^{J} g_j = \sum_{j=-\infty}^{J} \| g_j \|_{\mathrm{L}^2(\mathbb{R})} h_j, \qquad h_j = \frac{g_j}{\| g_j \|_{\mathrm{L}^2(\mathbb{R})}}.$$

This series converges since $\{h_j\}$ is orthonormal (since $\{g_j\}$ is orthogonal, as each g_j is in a different W_j) and eq 7.6 shows that $\sum \|g_j\|^2$ is bounded by $\|f\|^2$ and is thus finite. Also, all of the g_j's are in V_{J+1}, so any finite sum of them is in V_{J+1}, and this space is closed (by MRA assumption), so $g \in V_{J+1}$. Now note that $f - g$ can be written

$$\begin{aligned} f - g &= f_m + \sum_{j=m}^{J} g_j - \sum_{j=-\infty}^{J} g_j \\ &= f_m - \sum_{j=-\infty}^{m-1} g_j \end{aligned}$$

For every $m \leq J$, both terms on the R.H.S. are in V_m, so $f - g \in V_m$. We know that $\bigcap V_j = \{0\}$ from the MRA assumption, so $f = g$. Then, we see that $V_{J+1} \subseteq \mathcal{W}_J$ and thus $\mathcal{W}_J = V_{J+1}$

Part (i.) It is straight-forward to conclude that

$$\overline{\bigcup V_j} = \mathrm{L}^2(\mathbb{R}) \implies \overline{\bigoplus_{j=-\infty}^{\infty} W_j} = \mathrm{L}^2(\mathbb{R})$$

Part (j.) It remains only to show that $\{\psi_{m,n}\}_{m,n\in\mathbb{Z}}$ is an ONB for $\mathrm{L}^2(\mathbb{R})$. This follows immediately from the fact that $\{\psi_{m,n}\}_{n\in\mathbb{Z}}$ is an ONB for W_m and $W_m \perp W_n$ whenever $m \neq n$. That completes the result. □

Finally, Professor Benedetto mentioned that [7] discussed the interplay between the various MRA assumptions and cited several books and articles that included details on how to place these functions in various function spaces [3, 8, 9, 10, 11].

Haar MRA Discussion

Proposition 7.2 (Haar MRA)**.** *Let* $\varphi = \mathbb{1}_{[0,1)}$ *and let* $V_0 = \overline{span}\{\tau_n\varphi\}$. *Define* $V_j = \{f \in L^2(\mathbb{R}) \mid 2^{j/2} \cdot f(2^j t) \in V_0\}$. *Then,* $\{V_j\}, \varphi$ *is an MRA.*

Proof. By definition, the V_j are closed linear subspaces of $\mathrm{L}^2(\mathbb{R})$. That they union to all of $\mathrm{L}^2(\mathbb{R})$ and intersect only in the zero element follow from basic properties of integration theory. ◻

In this case, in the notation of the MRA theorem,

$$\begin{aligned} h_0[n] &= \frac{1}{\sqrt{2}} \cdot \int_{\mathbb{R}} \varphi(t/2) \cdot \varphi(t-n) dt \\ &= \frac{1}{\sqrt{2}} \cdot \int_0^2 \mathbb{1}_{[0,1)}(t-n) dt \\ &= \begin{cases} \frac{1}{\sqrt{2}} & \text{if } n = 0, 1 \\ 0 & \text{OW} \end{cases} \end{aligned}$$

Then, to determine ψ, we consider

$$h_1[n] = (-1)^n \cdot \overline{h_0[-n+1]} = \begin{cases} 1/\sqrt{2} & \text{if } n = 0 \\ -1/\sqrt{2} & \text{if } n = 1 \\ 0 & \text{OW} \end{cases} \implies \psi(t) = \sum h_1[n] \cdot \sqrt{2} \cdot \varphi(2t-n)$$

Note: This is practically the same as the $(-1)^{n+1}\overline{h_0[-n-1]}$ form developed in the theorem since neither integer translates nor multiplying by -1 affects the basis properties of the wavelet system.

We see fairly easily that $\psi(t) = \mathbb{1}_{[0,1/2)}(t) - \mathbb{1}_{[1/2,1)}(t)$, which is what we've defined as the Haar wavelet.

Lecture 8

28 October 2008

Separate beamer presentation provided by Emily King.

Lecture 9

04 November 2008

This begins with some references for frames, see [4] and [16].

Definition (MRA Frames)**.** The collection $\{V_j\}_{j\in\mathbb{Z}},\ \varphi$ is a *frame MRA* of $L^2(\mathbb{R})$ if

1. V_j is a closed linear subspace.
2. $V_j \subseteq V_{j+1}$.
3. $\bigcap V_j = \{0\}$.
4. $\overline{\bigcup V_j} = L^2(\mathbb{R})$.
5. $\{\tau_n\varphi\}$ is a frame for $\overline{\text{span}}\{\tau_n\varphi\} = V_0$.

There is a similar theory to MRAs, but we are most interested in the case of Riesz bases.

Proposition 9.1 (Haar MRA for $L^p(\mathbb{R})$, $p \in [0,\infty)$)**.** *Let* $\varphi = \mathbb{1}_{[0,1)}$ *and set*

$$V_0 = \left\{ \sum a_n \cdot \tau_n\varphi \mid \{a_n\} \in \ell^p(\mathbb{Z}) \right\}$$

Then, for each $j \in \mathbb{Z}$,

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

Then, $\{V_j\}$ *is a frame MRA of* $L^p(\mathbb{R})$ *with scaling function* φ*. If* $p = 2$*, then* $\{\tau_n\varphi\}$ *is an ONB for* V_0*. If* $p \neq 1, 2$*, then* $\{\tau_n\varphi\}$ *is a bounded, unconditional (Riesz) basis for* V_0*. It may not be the case when* $p = 1$ *(Prof. Benedetto unsure??), because no unconditional basis exists for* $L^1(\mathbb{R})$*, see [13].*

Proposition 9.2. *Suppose that* B *is a Banach space which has a Schauder basis, then* B *is separable.*

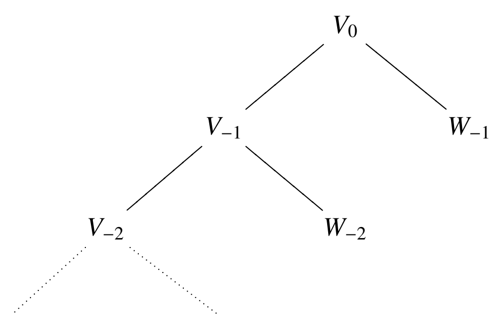
It is interesting that the converse of this proposition is not true, see [5] (this seems to be an interesting aside).

Now, suppose that $\{V_j\}$, φ is an MRA of $L^2(\mathbb{R})$. Then, we know (definition) that $\{\tau_n\varphi\}$ is an ONB for V_0, $V_0 = V_{-1} \oplus W_{-1}$, and $\{\tau_n\psi\}$ is an ONB for W_{-1}. Then, we see that

$$\left\{\frac{1}{\sqrt{2}} \cdot \varphi\Big(\frac{t}{2} - n\Big),\ \frac{1}{\sqrt{2}} \cdot \psi\Big(\frac{t}{2} - n\Big)\right\}$$

forms an ONB for V_0. So, we have constructed another ONB for V_0.

Recall that we have the tree decomposition



So that we have the decomposition

$$V_0 = V_{-m} \oplus \Big(\bigoplus_{i=0}^{m-1} W_{-m+i}\Big).$$

Then, the sequence of functions

$$\{\varphi_{-m,n}, \psi_{-m,n}, \psi_{-m+1,n}, \ldots, \psi_{-1,n} \mid n \in \mathbb{Z}\}$$

also forms an ONB for V_0. So, we naturally have a large class of distinct ONBs for V_0. This is a little strange, because our tree is *one-sided*. We will naturally extend this construction to a *full* dyadic tree to produce Wavelet packets.

First, a relevant aside. Let us define an ordering on the integers compatible with the dyadic structure of such a tree, called the *bit-reversal ordering*. We will do this inductively on the *level*, for the integers $\{0, 1, \ldots, 2^r - 1\}$, where r is the level. For level 1, we order $\{0, 1\}$ as the tuple $(0, 1)$. Not too interesting so far. For level 2, we order $\{0, 1, 2, 3\}$ as the tuple $(0, 2, 1, 3)$.

In general, if we've ordered $\{0, 1, \ldots, 2^m - 1\}$ as the tuple

$$(b_0, b_1, \ldots, b_{2^m-1}),$$

then we order $\{0, 1, \ldots, 2^{m+1} - 1\}$ as the tuple

$$(2b_0, 2b_1, \ldots, 2b_{2^m-1}, 2b_0 + 1, 2b_1 + 1, \ldots, 2b_{2^m-1} + 1)$$

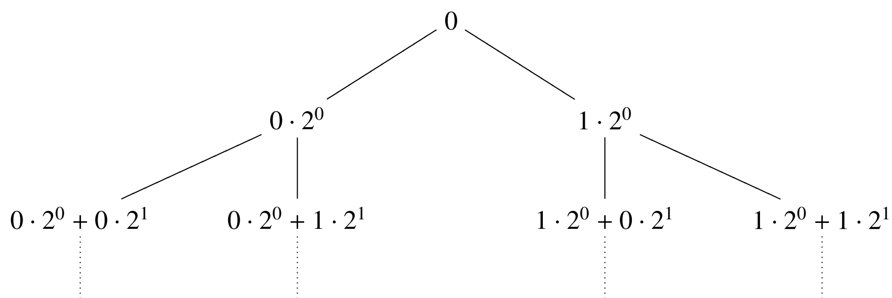
To motivate the name of this order, suppose we write the binary expansion for $n \in \mathbb{N}$, as

$$n = \sum_{j=1}^{\infty} \epsilon_j 2^{j-1}, \ \epsilon_j \in \{0, 1\}$$

If we consider the map $f: \mathbb{N} \to [0, 1)$ defined by

$$f(n) = \sum_{j=1}^{\infty} \epsilon_j 2^{-j},$$

which has the binary expansion for n written backwards (hence, bit-reversal). If we order $\mathbb{N}$ with the natural ordering of the image $f(\mathbb{N})$ inherited from the rational numbers, then we get the above defined order. To see this in a picture, we construct a tree as follows:



Now, returning to wavelet packets. From the MRA theorem, given $\{V_j\}, \varphi$ we construct ψ, H_0, H_1 s.t.

$$\sqrt{2} \cdot \hat{\varphi}(2\gamma) = \hat{\varphi}(\gamma) \cdot H_0(\gamma) \text{ and } \hat{\psi}(2\gamma) = \varphi(\gamma) \cdot H_1(\gamma)$$

Now, consider that for V_{-1}

$$\varphi_{-1,n}(t) = (1/\sqrt{2}) \cdot \varphi(t/2 - n) \implies \hat{\varphi}_{-1,n}(\gamma) = e_{-2n}(\gamma) \cdot (\sqrt{2}\hat{\varphi}(2\gamma)) \text{ and}$$
$$\sqrt{2} \cdot \hat{\varphi}(2\gamma) = \hat{\varphi}(\gamma) \cdot H_0(\gamma)$$

Similarly, for W_{-1}

$$\psi_{-1,n}(t) = (1/\sqrt{2}) \cdot \psi(t/2 - n) \implies \hat{\psi}_{-1,n}(\gamma) = e_{-2n}(\gamma) \cdot (\sqrt{2}\hat{\psi}(2\gamma)) \text{ and}$$
$$\sqrt{2} \cdot \hat{\psi}(2\gamma) = \hat{\varphi}(\gamma) \cdot H_1(\gamma)$$

Continuing to the next level, for V_{-2}

$$\varphi_{-2,n}(t) = (1/\sqrt{2})^2 \cdot \varphi(t/2^2 - n) \implies \hat{\varphi}_{-2,n}(\gamma) = e_{-4n}(\gamma) \cdot (\sqrt{2})^2 \hat{\varphi}(4\gamma) \text{ and}$$
$$(\sqrt{2})^2 \cdot \hat{\varphi}(4\gamma) = \hat{\varphi}(\gamma) \cdot H_0(\gamma) \cdot H_0(2\gamma)$$

and W_{-2}

$$\psi_{-2,n}(t) = (1/\sqrt{2})^2 \cdot \psi(t/2^2 - n) \implies \hat{\psi}_{-2,n}(\gamma) = e_{-4n}(\gamma) \cdot (\sqrt{2})^2 \hat{\psi}(4\gamma) \text{ and}$$
$$(\sqrt{2})^2 \cdot \hat{\psi}(4\gamma) = \hat{\varphi}(\gamma) \cdot H_0(\gamma) \cdot H_1(2\gamma)$$

Now, define

$$V_0 = X^0, V_{-1} = X_0^1, W_{-1} = X_1^1,\ V_{-2} = X_{(0,0)}^2, W_{-2} = X_{(0,1)}^2$$

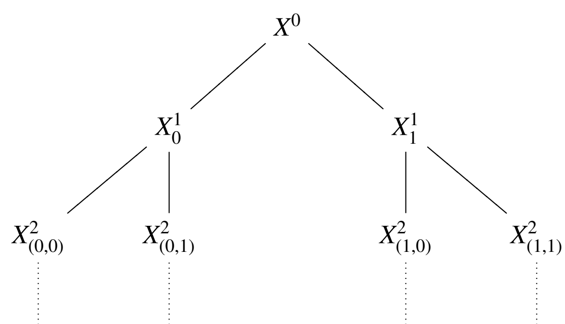
and $\theta_{(1,0)}, \theta_{(1,1)}$ by the formulas

$$(\sqrt{2})^2 \cdot \theta_{(1,0)}(\gamma) = \hat{\varphi}(\gamma) \cdot H_1(\gamma) \cdot H_0(2\gamma),\ (\sqrt{2})^2 \cdot \theta_{(1,1)} = \hat{\varphi}(\gamma) \cdot H_1(\gamma) \cdot H_1(2\gamma)$$

Now, define

$$X_{(1,0)}^2 = \overline{\text{span}}\{(1/\sqrt{2})^2 \cdot \theta_{(1,0)}(t/2^2 - n) \mid n \in \mathbb{Z}\}$$
$$X_{(1,1)}^2 = \overline{\text{span}}\{(1/\sqrt{2})^2 \cdot \theta_{(1,1)}(t/2^2 - n) \mid n \in \mathbb{Z}\}$$

Then, we have the tree



where each node is the direct sum of the two nodes below it. Of course, this can be continued inductively.

Proposition 9.3. *Assume that* $\{V_j\}, \varphi$ *is an MRA. Let* $m \in \{0, \dots, 2^r - 1\}$ *and*

$$m = \sum_{j=1}^{r} \epsilon_j \cdot 2^{j-1},\ \epsilon_j \in \{0, 1\}.$$

Define the function $\theta_{(\epsilon_1,\dots,\epsilon_r)}$ *by*

$$(\sqrt{2})^r \cdot \hat{\theta}_{(\epsilon_1,\dots,\epsilon_r)}(2^r\gamma) = \hat{\varphi}(\gamma) H_{\epsilon_1}(\gamma) \cdot H_{\epsilon_2}(2\gamma) \cdots H_{\epsilon_r}(2^{r-1}\gamma)$$

Let X_m^r *be defined by*

$$X_m^r = \overline{span}\{(1/\sqrt{2})^r \theta_{(\epsilon_1,\dots,\epsilon_r)}(t/2^r - n) \mid n \in \mathbb{Z}\}.$$

Then, if $k < r$

$$X_n^k = \bigoplus_{i \in I} X_i^r, \ I = \{i \mid 0 \leq i < 2^r, \ i \equiv n \ (2^k)\}$$

In particular, we have

$$V_0 = \bigoplus_{i=1}^{2^r-1} X_i^r$$

Lecture 10

11 November 2008

The lecture begins with a rehash of the last lecture, which I am omitting for brevity.

Definition. Let $H_0, H_1 \in \mathrm{L}^2(\mathbb{T})$ with Fourier coefficients $\{h_0[n]\}$ and $\{h_1[n]\}$ respectively. A sequence $\{w_i\}_{i\in\mathbb{N}} \subseteq \mathrm{L}^2(\mathbb{R})$ of functions is a sequence of *wavelet packets* corresponding to H_0, H_1, if

$$\begin{aligned} w_{2i}(t) &= \sum_{n\in\mathbb{Z}} h_0[n] \cdot \sqrt{2} \cdot w_i(2t-n) \text{ (in } \mathrm{L}^2(\mathbb{R})) \\ w_{2i+1}(t) &= \sum_{n\in\mathbb{Z}} h_1[n] \cdot \sqrt{2} \cdot w_i(2t-n) \text{ (in } \mathrm{L}^2(\mathbb{R})) \end{aligned}$$

Consider the Haar MRA. So, $\varphi = \mathbb{1}_{[0,1)}$ and H_0, H_1 s.t.

$$h_0[n] = \begin{cases} 1/\sqrt{2} & \text{if } n = 0,1 \\ 0 & \text{OW} \end{cases}$$
$$h_1[n] = \begin{cases} -1/\sqrt{2} & \text{if } n = 0,1 \\ 0 & \text{OW} \end{cases}$$

Therefore, we see that

$$\begin{aligned} H_0(\gamma) &= \frac{1}{\sqrt{2}} \cdot (1 + e^{-2\pi i\gamma}) = \sqrt{2} \cdot e^{-\pi i\gamma} \cdot \cos(\pi\gamma) \\ H_1(\gamma) &= \frac{1}{\sqrt{2}} \cdot (1 - e^{-2\pi i\gamma}) = \sqrt{2} \cdot e^{-\pi i\gamma} \cdot \sin(\gamma) \end{aligned}$$

Recall then that we compute $\psi = \mathbb{1}_{[0,1/2)} - \mathbb{1}_{[1/2,1)}$. Now, let $j = 0$, then

$$\begin{aligned} w_0(t) &= w_0(2t) + w_0(2t-1) \\ w_1(t) &= w_0(2t) - w_0(2t-1) \end{aligned} \tag{10.1}$$

Note that if $w_0 = \varphi$ and $w_1 = \psi$, then eq. 10.1 is fulfilled. *There is a pictorial representation here I am neglecting.*

Let's consider what other potential solutions to eq. 10.1 would entail. Note that by taking the Fourier transform of both sides (assuming that we can), we obtain:

$$\begin{aligned}
\hat{w}_0(\gamma) &= \frac{1}{2} \cdot (1 + e^{-pi\gamma}) \cdot \hat{w}_0\big(\gamma/2\big) \\
&= \frac{1}{\sqrt{2}} \cdot H_0\big(\gamma/2\big) \cdot \hat{w}_0\big(\gamma/2\big) \\
&= \frac{1}{\sqrt{2}} \cdot H_0\big(\gamma/2\big) \cdot \frac{1}{\sqrt{2}} \cdot H_0\big(\gamma/2^2\big) \cdot \hat{w}_0\big(\gamma/2^2\big) \\
&= \quad \vdots \qquad\qquad \vdots \qquad\qquad \vdots \\
&= \hat{w}_0\big(\gamma/2^n\big) \cdot \prod_{j=1}^{n} \Big(\frac{1}{\sqrt{2}} \cdot H_0\big(\gamma/2^j\big) \Big) \\
&= \hat{w}_0\big(\gamma/2^n\big) \cdot \prod_{j=1}^{n} \Big(e^{-\pi i(\gamma/2^j)} \cdot \cos\big(\pi\gamma/2^j\big) \Big)
\end{aligned}$$

Assuming that $w_0 \in \mathrm{L}^1(\mathbb{R})$ and $\hat{w}_0(0) = 1$, then

$$\hat{w}_0(\gamma) = \exp\Big(-\pi i\gamma \cdot \sum_{j=1}^{\infty} 2^{-j} \Big) \cdot \prod_{j=1}^{\infty} \cos\big(\pi\gamma/2^j\big) = e^{-\pi i\gamma} \cdot \frac{\sin(\pi\gamma)}{\pi\gamma}$$

Then, we get $w_0 = \varphi$ and $w_1 = \psi$.

Definition. Define the sequence of functions $\{R_n\}_{n\in\mathbb{N}}$ defined by

$$R_n(t) = \operatorname{sgn}\big(\sin\big(2^n \pi t\big)\big)$$

are called the *Rademacher functions.*

Proposition 10.1. *The sequence* $\{R_n\}_{n\in\mathbb{N}}$ *is orthonormal.*

Proposition 10.2. *The sequence* $\{R_n\}_{n\in\mathbb{N}}$ *does not form an orthonormal basis for* $L^2(\mathbb{T})$*. In fact, note that*

$$\int \cos(2\pi t) \cdot R_n(t)\, dt, \qquad \forall n \in \mathbb{N}.$$

Definition. The sequence of functions $\{w_j\}$ defined by

$$n = \sum_{j=1}^{\infty} \varepsilon_j \cdot 2^{j-1}, \;\; w_n(t) = \prod_{\varepsilon_j = 1} R_j(t)$$

are called the *Walsh functions*.

Theorem 10.3 (Walsh-1923)**.** *The sequence* $\{w_n\}_{n\in\mathbb{N}}$ *forms an ONB for* $L^2(\mathbb{T})$.

Proof. The orthonormality follows from the orthonormality of the Rademacher functions. It suffices to prove that

$$\overline{\text{span}}\{w_n\} = \mathrm{L}^2(\mathbb{T})$$

Pick $f \in \mathrm{L}^2(\mathbb{T})$, and suppose that

$$\int_0^1 f(t) \cdot w_n(t)\, dt = 0, \ \forall n \in \mathbb{N}$$

We know that $\mathrm{L}^1(\mathbb{T}) \subset \mathrm{L}^2(\mathbb{T})$. Then, the function

$$g(t) = \int_0^t f(s)\, ds$$

is well-defined. It is continuous, by elementary properties of integration. Note that

$$0 = \int_0^1 f(t) w_1(t)\, dt = \int_0^{1/2} f(t)\, dt - \int_{1/2}^1 f(t)\, dt = 2 \cdot g(1/2)$$

This process can be continued iteratively accross the w_j to conclude that $g(t) = 0$ whenever t is a dyadic rational number. We know that g is continuous and the dyadic rationals are a dense subset of $[0, 1)$, so g must be identically 0. Then, f is identically 0, and we're done. □

Lecture 11

18 November 2008

11.1 Compressed Sensing Introduction

Definition. *Compressed sensing* is a sensing or sampling theory attempting to recover certain signals from far fewer samples or measurements than traditionally used. There are two fundamental concepts underlying compressed sensing:

1. Sparsity
2. Incoherence

11.1.1 Setup

We shall consider discrete time signals

$$f : \mathbb{Z}_n \to \mathbb{R}$$

You can think of such a signal as a vector $f \in \mathbb{R}^n$. We are interested in sensing mechanisms in which information about f is obtained through linear functionals

$$y_k = \langle f, \varphi_k \rangle, \ k = 1, \ldots, m, \ \text{where } n \gg m$$

where m is the number of samples, and we have undersampled.

Example 11.1 (Spikes or Delta Functions)**.** Consider $\varphi_k(t) = \delta(t-k)$, so

$$y_k = f(k) = \sum_{t=0}^{n-1} f(t) \cdot \delta(t-k)$$

Example 11.2 (Sinusoids)**.** Consider $\varphi_k(t) = (1/\sqrt{n}) \cdot e^{-2\pi i k t/n}$, and

$$y_k = \hat{f}(k) = \frac{1}{\sqrt{n}} \cdot \sum_{t=0}^{n-1} f(t) \cdot e^{-2\pi i k t/n}$$

11.1.2 Shannon Sampling Theory

Suppose that $f \in \mathbb{R}^n$ and suppose that

$$\text{supp}(\hat{f}) \subseteq \{\omega_1, \dots, \omega_\ell\} \subseteq \hat{\mathbb{Z}}_n, \text{ where } \omega_1 < \omega_2 < \cdots < \omega_\ell$$

Shannon Sampling tells us that in order to generate a perfect reconstruction of f, then we need $m = \omega_\ell - \omega_1$ equally spaced samples in the time domain.

Note that if $\omega_\ell - \omega_1$ is large, then we need to take a large number of measurements to reconstruct f, even if $|\text{supp}(\hat{f})|$ is small.

11.1.3 Sparsity

Let $f \in \mathbb{R}^n$, and $\Psi = [\psi_1, \dots, \psi_n]$ be an ONB for $\mathbb{R}^n$. We can represent f in terms of Ψ as follows:

$$f(t) = \sum_{k=0}^{n-1} x_k \psi_k(t), \; t \in \mathbb{R}^n$$
$$f(t) = \Psi \cdot x$$

Definition. We say that f is *S-sparse* in Ψ, if all but S of the x_k are 0. Similarly, f is *S-compressible* on Ψ if the x_k are concentrated on a set of size S.

Given $x = \{x_k\}$, let x_S denote the vector defined by setting all but the largest S coefficients of x to 0. Then,

$$x \text{ is S-sparse} \iff \| x - x_S \|_2 = 0$$
$$x \text{ is S-compressible} \iff \| x - x_S \|_2 < \varepsilon$$

Then, if $f = \Psi \cdot x$, then $f_S = \Psi \cdot x_S$ is a good approximation to f. The drawback of this approach is that it requires sorting and processing all n coefficients of f before reducing to f_S.

11.1.4 Incoherent Sampling

Let (Φ, Ψ) be a pair of ONB for $\mathbb{R}^n$. We can use Φ to sense the signal, and Ψ will be used to represent the signal f.

Definition. The *coherence* between Φ and Ψ is defined as

$$\mu(\Phi, \Psi) = \sqrt{n} \cdot \max_{1 \leq j,k \leq n} \left| \langle \varphi_k, \psi_j \rangle \right|,$$

which measures the largest correlation between pairs of elements of Φ and Ψ.

Clearly, a large value of coherence means that Φ and Ψ contain correlated vectors, and a small value of coherence means that they do not. The values for coherence are actually bounded.

Proposition 11.3. $1 \leq \mu(\Phi, \Psi) \leq \sqrt{n}$

The upper bound is clear, given the assumption of orthonormality. The lower bound is obtained as a limit letting Φ be a spike basis, and Ψ be a sinusoidal basis.

11.1.5 Undersampling and Sparse Signal Recovery

[Undersampling]
Let $M \subset \mathbb{Z}_n^*$ with $|M| = m$, where M is the index set of our sampling coefficients:

$$y_k = \langle f, \varphi_k \rangle, \ k \in M, \ \varphi \in \Phi = \text{ONB for } \mathbb{R}^n$$

Our goal is to reconstruct f from our samples y_k.

Compressed Sensing proposes the reconstructed signal $f^* = \Psi \cdot x^*$ as its best guess for f, where x^* is the solution to the following convex optimization problem:

$$x^* = \min_{\tilde{x} \in \mathbb{R}^n} \| \tilde{x} \|_1, \ \text{s.t. } y_k = \langle \Psi_{\tilde{x}}, \varphi_k \rangle \text{ and } \| \tilde{x} \|_1 = \sum_{k=0}^{n-1} |\tilde{x}_k| \tag{11.1}$$

So, x^* is the coefficient sequence with minimal ℓ^1 norm, consistent with our sampled data y_k.

Proposition 11.4. *Fix $f \in \mathbb{R}^n$ and suppose that the coefficient sequence x of f in the ONB Ψ is S-sparse. Select m measurements in the Φ domain uniformly at random. Then, if*

$$m \geq C \cdot \left(\mu(\Phi, \Psi)\right)^2 \cdot S \cdot \log(n)$$

for some constant C, the solution to eq. 11.1 is exact (i.e. $f^* = \Psi \cdot x^* = f$) with (high) probability related to the value of C.

There are several important things to note about this result:

1. The role of μ, the coherence between Φ and Ψ. The more incoherent (the less related) that Φ and Ψ, the fewer measurements are necessary.

2. We are sampling non-adaptively (randomly, not intelligently).

3. One suffers no information loss by measuring just about any set of m coefficients, which may be far less than the signal size demands according to Shannon sampling theory. If all we know is that $\text{supp}(\hat{f}) \subset \mathbb{Z}_n$, then Shannon sampling theory asserts we need $m = n$ samples, but the CS result indicates we need only $m \geq C \cdot S \cdot \log(n)$ samples.

4. The decoder eq. 11.1 assumes no knowledge of x.

5. The role of probability: there exist certain outstanding signals that vanish a.e. in the Φ domain.
   For example, let $n = N^2$ and set

   $$\Phi = \text{sinusoids}, \ \Psi = \text{spikes}, \ f(t) = \begin{cases} 1 & \text{if } t = k \cdot N, \ k = 0, 1, \ldots, N-1 \\ 0 & \text{OW} \end{cases}$$

   It can be shown that $f = \hat{f} \to N^2 - N$, and the fourier coefficients of f are 0 most of the time. Random samples will be mostly 0, contrary to the result.

11.2 Restricted Isometry Problem

11.2.1 Setup

We want to recover a signal $f \in \mathbb{R}^\ell$ from corrupted measurements $y = A \cdot f + e$, where A is an $n \times \ell$ coding matrix (full rank, $n > \ell$), and e is an unknown vector of errors.

Suppose that

$$\| e \|_0 = |\{k \mid e_k \neq 0\}| = S$$

For what values of S can we recover f from the measurement y? Notice that to reconstruct f, it suffices to determine e, since $y - e = A \cdot f$ and A has full rank.

We want to construct an $m \times n$ matrix Φ s.t. $\Phi \cdot A = 0$ (take Φ s.t. $\ker(\Phi) = \text{range}(A)$), then apply Φ to our measurements y:

$$\tilde{y} = \Phi \cdot y = \Phi \cdot (A \cdot f + e) = \Phi \cdot e$$

In order to find f, we want to reconstruct an S-sparse vector e from measurements $\tilde{y} = \Phi \cdot e$. Note that $\tilde{y}_k = \langle e, \varphi_k \rangle$, where $\overline{\varphi_k}$ is the k^{th} row of Φ.

If e is sparse enough, then it should be the solution to

$$\min_{d \in \mathbb{R}^n} \| d \|_0 \text{ subject to } \Phi \cdot d = \tilde{y}\, (= \Phi \cdot e) \tag{11.2}$$

Note it is impossible to solve eq. 11.2 for even modestly sized signals. As an alternative, consider

$$\min_{d \in \mathbb{R}^n} \| d \|_1 \text{ s.t. } \Phi \cdot d = \tilde{y} \tag{11.3}$$

11.2.2 Restricted Isometries

Let $(v_j)_{j \in J} \in \mathbb{R}^m$ denote the columns of Φ. For any subset $T \subseteq J$, define Φ_T to be the submatrix with column choices $j \in T$ s.t.

$$\Phi_T \cdot C = \sum_{j \in T} c_j \cdot v_j$$

Definition. Let Φ be the matrix with the finite collection of vectors $(v_j)_{j \in J} \in \mathbb{R}^m$ as columns. For every integer $1 \leq S \leq |J|$, we define the *S-restricted isometry constant* δ_S to be the smallest quantity s.t. Φ_T obeys

$$(1 - \delta_S) \cdot \| C \|_2 \leq \| \Phi_T \cdot C \|_2 \leq (1 + \delta_S) \cdot \| C \|_2$$

for all subsets $T \subseteq J$ of cardinality at most S.

Any matrix Φ satisfying the above definition for some S and $1 < \delta_S < 1$ is said to satisfy the *restricted isometry property (RIP)* of order S.

Lemma 11.5. *Suppose that $S \geq 1$ s.t. $\delta_{2S} < 1$ and let $T \subset J$ with $|T| \leq S$. If $f = \Phi_T \cdot C$ for some arbitrary $|T|$-dimensional vector C, then the set T and the coefficients $(c_j)_{j \in T}$ can be reconstructed uniquely from knowledge of the vector f and v_j's alone (i.e. eq. 11.2 has a unique solution).*

Proof. It suffices to prove that C is unique. Suppose that

$$f = \Phi_T \cdot C = \Phi_{T'} \cdot C'$$

By assumption, $|T|, |T'| \leq S$. Consider the vector $d = (d_j)$ defined by

$$d_j = c_j - c'_j$$

By assumption, $\Phi_{T \cup T'} \cdot d = 0$ and Φ satisfies the RIP of order $2S$. Then,

$$(1 - delta_{2S}) \cdot \| d \|_2 \leq \| \Phi_{T \cup T'} \cdot d \|_2 \leq (1 + \delta_{2S}) \cdot \| d \|_2$$

The middle term is 0, and $(1 - \delta_{2S}), (1 + \delta_{2S}) > 0$, so we see that d must be the zero vector. In other words, $C = C'$, as required. □

Proposition 11.6. *Suppose that $S \geq 1$ s.t. $\delta_S + \delta_{2S} + \delta_{3S} < 1$ and let C be a real vector supported on a set $T \subseteq J$ s.t. $|T| \leq S$. Set $f = Phi \cdot C$, then C is the unique minimizer to eq: 11.3 where*

$$(1 - \delta_S) \cdot \| C \|_2 \leq \| \Phi_T \cdot C \|_2 \leq (1 + \delta_S) \cdot \| C \|_2$$

Returning to our original problem of reconstructing f from $y = A \cdot f + e$, the above proposition states that if e is S-sparse, and Φ has $\delta_S + \delta_{2S} + \delta_{3S} < 1$, then one can recover e, and hence f, through solving eq. 11.3.

11.2.3 Good RIP matrices

The following two types of $m \times n$ matrices satisfy the RIP of order S with high probability provided that $m \geq C \cdot S \cdot \log(n/S)$:

1. Matrices with i.i.d. entries following a Gaussian distribution with mean 0 and variance $1/m$.
2. Matrices with i.i.d. entries following a Bernoulli distribution with mean 0 and variance $1/m$.

This lower bound is near optimal.

Lecture 12

25 November 2008

Proposition 12.1 (Gauss)**.** *Let G be a finite, abelian group. Then, G is isomorphic to*

$$\prod_{i=1}^{n} \mathbb{Z}_{N_i}, \text{ where } N_i = p_i^{k_i}, \ p_i \text{ prime}$$

Proof of this may be found in most algebra books under the classification of finitely generated modules over $\mathbb{Z}$ (or a PID).

Example 12.2. Let N be an integer. Then, consider $\mathbb{Z}_N = \mathbb{Z}/n\mathbb{Z}$. If N is prime, then $\mathbb{Z}_N$ is a finite field. Otherwise, $\mathbb{Z}_N$ is a commutative ring with unity, which has (the obvious) zero-divisors.

Definition. Let G be a finite group. The *characters* of G are the homomorphisms

$$e\colon G \to \mathbb{C}$$

Note that it is implied that $|e(g)| = 1$. These characters form a group themselves under the operation of multiplication. Denote the *group of characters* of G as $\hat{G}$.

Proposition 12.3. *For G a finite abelian group, the group $\hat{G}$ is isomorphic to G.*

Definition. The *Discrete Fourier Transform* $\hat{x}$ of $x\colon G \to \mathbb{C}$ (not necessarily a character) is defined as

$$\forall e \in \hat{G}, \ \hat{x}(e) = \sum_{g \in G} x(g) \cdot \overline{e(g)}$$

Define an inner product on $\ell^2(G)$ by

$$x, y \in \ell^2(G), \ \langle x, y \rangle = \sum_{g \in G} x(g) \cdot \overline{y(g)}$$

Example 12.4. Take $G = \mathbb{Z}_N$, $x\colon \mathbb{Z}_N \to \mathbb{C}$, then

$$\hat{x}[n] = \sum_{m=0}^{N-1} x[m] \cdot e^{-2\pi imn/N}$$

Proposition 12.5. *Let G be a finite abelian group.*

1. $\forall e_1, e_2 \in \hat{G}$, *then*

$$\langle e_1, e_2 \rangle = \begin{cases} |\hat{G}| = |G| & \text{if } e_1 = e_2 \\ 0 & \text{otherwise} \end{cases}$$

2. $\forall g_1, g_2 \in G$, *then*

$$\sum_{e \in \hat{G}} e(g_1) \cdot \overline{e(g_2)} = \begin{cases} |\hat{G}| = |G| & \text{if } g_1 = g_2 \\ 0 & \text{otherwise} \end{cases}$$

Theorem 12.6 (Inversion & Plancherel)**.** *Let G be a finite abelian group.*

1. $\forall x \in \ell^2(G),\ \forall g \in G$,

$$x(g) = \frac{1}{|\hat{G}|} \cdot \sum_{e \in \hat{G}} \hat{x}(e) \cdot e(g)$$

2. $\forall x \in \ell^2(G)$,

$$\| x \|_2 \stackrel{def}{=} \Big(\sum_{g \in G} |x(g)|^2 \Big)^{1/2} = \frac{1}{\sqrt{|\hat{G}|}} \cdot \Big(\sum_{e \in \hat{G}} |\hat{x}(e)|^2 \Big)^{1/2}$$

Proof. **Part (1.)** Consider

$$\begin{aligned} \sum_{e \in \hat{G}} \hat{x}(e) \cdot e(g_0) &= \sum_{e \in \hat{G}} \sum_{g \in G} x(g) \cdot \overline{e(g)} \cdot e(g_0) \\ &= \sum_{g \in G} x(g) \cdot \sum_{e \in \hat{G}} e(g_0) \cdot \overline{e(g)} \\ &= |G| \cdot x(g_0) \end{aligned}$$

Part (2.) Consider

$$\begin{aligned} \| x \|_2 &= \frac{1}{|\hat{G}|} \cdot \Big(\sum_{g \in G} |\hat{x}(e) \cdot e(g)|^2 \Big)^{1/2} \\ &= \frac{1}{|\hat{G}|} \cdot \Big(\sum_{g \in G} \sum_{e \in \hat{G}} \sum_{e' \in \hat{G}} \hat{x}(e) \cdot e(g) \cdot \overline{\hat{x}(e') \cdot e'(g)} \Big)^{1/2} \\ &= \frac{1}{|\hat{G}|} \cdot \Big(\sum_{e \in \hat{G}} \sum_{e' \in \hat{G}} \hat{x}(e) \cdot \overline{\hat{x}(e')} \cdot \sum_{g \in G} e(g) \cdot \overline{e'(g)} \Big)^{1/2} \\ &= \frac{1}{|\hat{G}|} \cdot \Big(\sum_{e \in \hat{G}} |G| \cdot |\hat{x}(e)|^2 \Big)^{1/2} \end{aligned}$$

□

Proposition 12.7 (Poisson Summation Formula)**.** *Let $x \in \ell^2(G)$, $H \trianglelefteq G$ (subgroup), and consider $(G/H)\hat{}$:*

$$\sum_{h \in H} x(h) = \sum_{e \in (G/H)\hat{}} \hat{x}(e)$$

Example 12.8 (More General)**.** Let G be a locally compact abelian group. The group $\hat{G}$ of (continuous) characters of G is still well-defined. In this context, the group $\hat{\hat{G}}$ is naturally isomorphic to G.

Example 12.9 (Dyadic Example)**.** Let

$$W_n = \prod_{i=1}^{n} \mathbb{Z}_2$$

and consider $\iota\colon W_n \to Q_n = \{0, 1, \ldots, 2^n - 1\}$ defined by

$$(\varepsilon_1, \ldots, \varepsilon_n) \mapsto \sum_{j=1}^{n} \varepsilon_j \cdot 2^{j-1}$$

This is the bijection defined by the dyadic expansion. By definition, ι carries the group operation from W_n to the set Q_n, but this is *not* addition $\mod (2^n)$.

Example 12.10 (Cantor Group)**.** Endow the group $\mathbb{Z}_2$ with the discrete topology, and define the topological group

$$\mathbb{D} = \prod_{i=1}^{\infty} \mathbb{Z}_2$$

With the product group operation (component-wise) and topological structure (product topology). Then, $\mathbb{D}$ is a totally disconnected, compact, abelian group called the *Cantor Group*. Note that $\mathbb{D}$ as a locally compact group, is endowed with a unique, invariant, Haar measure.

Proposition 12.11 (Fine 1949)**.** *$\hat{\mathbb{D}}$ is equivalent to the set of Walsh functions on $\mathbb{T}$.*

Proposition 12.12. *The map $\Phi\colon \mathbb{D} \to [0, 1)$ defined by*

$$(\varepsilon_j)_{j=1}^{\infty} \mapsto \sum_{j=1}^{\infty} \varepsilon_j \cdot 2^{-j}$$

maps the Haar measure on $\mathbb{D}$ to Lebesgue measure on $[0, 1)$.

Proposition 12.13. *For each $n \in \mathbb{N}$, there is a unique $\gamma_n \in \hat{\mathbb{D}}$ that is nontrivial* only *on the n^{th} factor. Actually, $\gamma_n(x) = e^{-\pi i(\pi_n(x))}$, where π_n is the canonical projection onto the n^{th} factor. Then, $\hat{\mathbb{D}} = \mathbb{W}$ is the set of finite products of the γ_n.*

Theorem 12.14. *In general, let G be a compact abelian group, endowed with Haar measure. Then, $\hat{G}$ is an ONB for $L^2(G)$.*

Proposition 12.15 (Kolomogorov 1922)**.** *Let $F \in L^2(\mathbb{T})$, then*

$$\lim_{n\to\infty} S_{2^n}(F)(x) = F(x) \text{ a.e., where } S_N(F)(x) = \sum_{n=-N}^{N} \hat{F}(n) \cdot e^{2\pi i n x}$$

I missed the relevance of this result at this particular place in the lecture. I believe he intended to connect the dots, but didn't have time and just continued.

Now, we return to the issue of generalizing to multiple dimensions.

Proposition 12.16. *Let $K \subseteq \mathbb{R}^d$ be Lebesgue measurable, so that K is τ-congruent to $[-1/2, 1/2]^d$ and the set*

$$\{2^m \cdot K \mid m \in \mathbb{Z}\}$$

is a tiling of $\mathbb{R}^d$. Let $\hat{\psi} = \mathbb{1}_K$. Then, $\{\psi_{m,n} \mid m \in \mathbb{Z}, n \in \mathbb{Z}^d\}$ is an orthonormal basis for $L^2(\mathbb{R}^d)$.

Proposition 12.17. *Assume that K_0 is τ-congruent to $[-1/2, 1/2]^d$, and*

$$T : K_0 \to [-2N, 2N]^d \setminus [-N, N]^d$$

be a measurable, injective function with the property that

$$\forall \gamma \in K_0,\ \exists k_\gamma \in \mathbb{Z}^d \text{ s.t. } T(\gamma) = \gamma + k_\gamma$$

Then, T is measure preserving.

Given K_0 and T as above, consider the following construction:
First, define

$$A_0 = K_0 \cap \Big(\bigcup_{j\geq 1} 2^{-j} \cdot K_0\Big),\ K_1 = (K_0 \setminus A_0) \cup T(A_0)$$

Then, $K_0 \setminus A_0 \subseteq K_0$ and $T(A_0) \subseteq [-2N, 2N]^d \setminus [-N, N]^d$.

Now, define

$$A_1 = K_1 \cap \Big(\bigcup_{j\geq 1} 2^{-j} \cdot K_1\Big),\ K_2 = ((K_0 \setminus A_0) \setminus A_1) \cup (T(A_0) \cup T(A_1)).$$

Then, $((K_0 \setminus A_0) \setminus A_1) \subseteq K_0$ and $T(A_0) \cup T(A_1) \subseteq [-2N, 2N]^d \setminus [-N, N]^d$.

Generally, define

$$A_n = K_n \cap \Big(\bigcup_{j\geq 1} 2^{-j} K_n\Big)$$

$$K_{n+1} = (K_0 \setminus A_0 \setminus A_1 \setminus \cdots \setminus A_n) \cup \Big(\bigcup_{k=0}^{n} T(A_k)\Big)$$

$$K = \Big(\bigcap_{n\in\mathbb{N}} (K_0 \setminus A_0 \setminus \cdots \setminus A_n)\Big) \cup \Big(\bigcup_{n\in\mathbb{N}} T(A_n)\Big)$$

This is the construction of a K which fulfills the requirements of proposition 12.16. Computationally, this construction may not be overly helpful, but it need not be carried to completion to be helpful.

Proposition 12.18. *A defined in the construction, for $n > 0$, $K_n \setminus A_n$ defines a parseval wavelet set, and K_n defines a frame wavelet set with frame bounds between* 1 *and* 2*.*

Lecture 13

01 December 2008

Let ψ be an orthonormal wavelet for $L^2(\mathbb{R})$. Now, $\forall j, k, m, n \in \mathbb{Z},\ (x_1, x_2) \in \mathbb{R}^2$, define

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

Theorem 13.1. *The set $\{\psi_{j,k} \otimes \psi_{m,n} \mid i, j, m, n \in \mathbb{Z}\}$ is an orthonormal basis for $L^2(\mathbb{R}^2)$. The terminology is that $\psi \otimes \psi$ is a* rectilinear orthonormal wavelet.

Remark: The set above consists of all possible dilates and translates of Ψ in each dimension. This is a gigantic set.

Let's now consider MRAs. Let $\{V_j\}$ be an MRA for $L^2(\mathbb{R})$ with scaling function φ. Define $\varphi \otimes \varphi(x_1, x_2) = \varphi(x_1) \cdot \varphi(x_2)$ and

$$V_0^2 = \{f : \mathbb{R}^2 \to \mathbb{C} \mid f(x_1, x_2) = \sum_{n_1, n_2 \in \mathbb{Z}} c_{n_1,n_2} \cdot \varphi \otimes \varphi(x_1 - n_1, x_2 - n_2) \text{ where } \{c_{n_1,n_2}\} \in \ell^2(\mathbb{Z}^2)$$

Proposition 13.2.

$$V_0^2 = \overline{span}\{f \otimes g : f, g \in V_0\} = \underbrace{V_0 \hat{\otimes} V_0}_{\text{projective tensor product}}$$

Similarly, define

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

Proposition 13.3. *$\{V_j^2\}_{j \in \mathbb{Z}}$ is an MRA for $L^2(\mathbb{R}^2)$ with scaling function $\varphi \otimes \varphi$ and orthonormal wavelets $\varphi \otimes \psi, \psi \otimes \varphi,\ \psi \otimes \psi$.*

Proof.

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

□

Definition. Let $W = \{\psi_1, \ldots, \psi_L\} \subset \mathrm{L}^2(\mathbb{R}^d)$. Then, W is a *set of dyadic wavelets* (for $\mathrm{L}^2(\mathbb{R}^d)$) if the set

$$\{(\psi_\ell)_{m,n} \mid m \in \mathbb{Z},\ n \in \mathbb{Z}^d,\ \ell = 1, \ldots, L\}$$

is an orthonormal basis for $\mathrm{L}^2(\mathbb{R}^d)$, where

$$(\psi_\ell)_{m,n}(x) = 2^{md/2} \cdot \psi_\ell(2^m \cdot x - n)$$

The set W is *associated with an MRA* $\{V_j\}$ with scaling function $\varphi \in \mathrm{L}^2(\mathbb{R}^d)$ if the set

$$\{\tau_n \psi_\ell \mid n \in \mathbb{Z}^d,\ \ell = 1, \ldots, L\}$$

is an ONB for W_0 where $V_1 = V_0 \oplus W_0 \subseteq \mathrm{L}^2(\mathbb{R}^d)$.

Proposition 13.4 (P. Auscher)**.** *If W is a set of dyadic wavelets s.t. each $\hat{\psi}_\ell$ has weak smoothness and decay (e.g. $|\hat{\psi}_\ell|$ is continuous and is $O(|\gamma|^{-(d/2+varepsilon)})$), then W is associated with an MRA.*

Proposition 13.5 (Auscher, Gripenberg, Wang)**.** *If W is a set of wavelets associated with an MRA in $L^2(\mathbb{R}^d)$, then $|W| = 2^d - 1$.*

Proposition 13.6. *The set $\{\psi_{m,n} \mid m \in \mathbb{Z},\ n \in \mathbb{Z}^d\} \subseteq L^2(\mathbb{R}^d)$ is orthonormal if and only if*

$$\sum_{n \in \mathbb{Z}^d} |\hat{\psi}(\gamma + n)|^2 = 1 \text{ a.e., and}$$

$$\forall m \geq 1,\ \sum_{n \in \mathbb{Z}^d} \hat{\psi}(2^m(\gamma + n)) \cdot \overline{\hat{\psi}(\gamma + n)} = 0 \text{ a.e.}$$

Proposition 13.7. *The set $\{\psi_\ell : \|\psi_\ell\|_{L^2(\mathbb{R}^d)} = 1,\ \ell = 1, \ldots, L\}$ is a dyadic set of wavelets for $L^2(\mathbb{R}^d)$ if and only if*

$$\sum_{\ell=1}^{L} \sum_{m \in \mathbb{Z}} |\hat{\psi}_\ell(2^m \cdot \gamma)|^2 = 1 \text{ a.e., and}$$

$$\forall n \in \mathbb{Z}^d \setminus (2\mathbb{Z}^d),\ \sum_{\ell=1}^{L} \sum_{m=0}^{\infty} \hat{\psi}_\ell(2^m \cdot \gamma) \cdot \overline{\hat{\psi}_\ell(2^m(\gamma + n))} = 0 \text{ a.e.}$$

Lecture 14

09 December 2008

Definition. Suppose that we have an invertible linear transformation

$$L\colon \mathbb{R}^d \to \mathbb{R}^d.$$

Then, we $L(\mathbb{Z}^d)$, a free $\mathbb{Z}$ module of rank d embedded in $\mathbb{R}^d$. We define a *lattice* Λ as the image $\mathbb{Z}^d$ under such an invertible linear transformation. The linear transformation will not be unique, but there will always be one which determines Λ as it's image. Note that it is the embedding in $\mathbb{R}^d$ that makes the lattice an object of interest. All lattices contained in $\mathbb{R}^d$ are isomorphic as $\mathbb{Z}$ modules (free, of rank d), it is the embedding which sets them apart.

Definition. Let $\Lambda \subset \mathbb{R}^d$ be a lattice, and let A be a $d \times d$ matrix with the following properties:

1. $A(\Lambda) \subseteq \Lambda$ (A leaves Λ invariant)
2. If λ is an eigenvalue of A, then $|\lambda| > 1$.

Then, A is called a *dilation matrix*. This is a generalization from the 1-dimensional dyadic case.

Example 14.1. Let $\Lambda = Z^d \subset \mathbb{R}^d$ and $A = m \cdot I_d$ for any integer $m > 1$, where I_d is the $d \times d$ identity matrix.

Lemma 14.2. *Let A be a dilation matrix for some lattice $\Lambda \subset \mathbb{R}^d$. Then,* $\det(A) \in \mathbb{Z}$*, and* $|\det(A)| > 1$.

Proof. Let $\{\lambda_1, \dots, \lambda_d\} \subset \mathbb{C}$ be the eigenvalues of A, with repetition as necessary. By assumption, $\forall i \in \{1, \dots, d\},\ |\lambda_i| > 1$. Then,

$$|\det(A)| = \left(\prod_{i=1}^{d} |\lambda_i| \right) > 1$$

It suffices to prove that the determinant is an integer. We know that Λ, so there exists an invertible linear transformation $L\colon \mathbb{R}^d \to \mathbb{R}^d$ s.t. $\Lambda = L(\mathbb{Z}^d)$. Also, A is a dilation matrix, so

$$A(L(\mathbb{Z}^d)) \subset L(\mathbb{Z}^d).$$

We know L is invertible, so we see

$$L^{-1} \circ A \circ L(\mathbb{Z}^d) \subset \mathbb{Z}^d$$

Then, the linear transformation $L^{-1} \circ A \circ L$ must have only integer entries, and $\det(L^{-1} \circ A \circ L)$ is a polynomial function of these integer entries. Then,

$$\det(L^{-1} \circ A \circ L) \in \mathbb{Z} \text{ and } \det(L^{-1} \circ A \circ L) = \det(A)$$

Then, we see that $\det(A) \in \mathbb{Z}$. □

Proposition 14.3. *Here a few relevant change of variables formulae:*

1. *Let $f \in L^1_{loc}(\mathbb{R}^d)$, $g \in L^\infty(\mathbb{R}^d)$ with compact support, and A be a linear automorphism of $\mathbb{R}^d$. Then,* $\det(A) \neq 0$ *and*

$$(f \circ A)(g) \stackrel{def}{=} \int_{\mathbb{R}^d} f(A(x)) \cdot g(x)\, dx = \frac{1}{|\det(A)|} \cdot \int_{\mathbb{R}^d} f(x) \cdot g(A^{-1}(x))\, dx$$

2. *Let A be a linear automorphism of $\mathbb{R}^d$, with adjoint A^*. Then, A^* is an automorphism of $\mathbb{R}^d$, $(A^*)^{-1} = (A^{-1})^a st$, and* $\det(A^*) = \overline{\det(A)}$.

3. *Let A be a linear automorphism of $\mathbb{R}^d$. If $f \in L^1(\mathbb{R}^d)$, then*

$$(f \circ A)\hat{} = \frac{1}{\det(A)} \cdot \hat{f} \circ (A^{-1})^*$$

Proposition 14.4. *Let A be a linear automorphism of $\mathbb{R}^d$. then, A induces a unitary operator*

$$\mathcal{U}_A : L^2(\mathbb{R}^d) \to L^2(\mathbb{R}^d) \text{ given by } f \mapsto |\det(A)|^{1/2} \cdot f \circ A$$

Definition. This extends the definition of an MRA given for 1-dimension. An MRA associated with a dilation matrix and lattice (A, Λ) is a scaling function $\varphi \in \mathrm{L}^2(\mathbb{R}^d)$ and "scaffolding" $\{V_j\}_{j\in\mathbb{Z}}$ such that the following hold:

1. $\forall j \in \mathbb{Z},\ V_j \subseteq V_{j+1}$.
2. $f \in V_0 \iff \tau_\lambda f \in V_0,\ \forall \lambda \in \Lambda$.
3. $f \in V_j \iff f \circ A(x) \in V_{j+1},\ \forall j \in \mathbb{Z}$ (i.e. $V_j = \mathcal{U}_A^j(V_0)$)
4. $$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = \mathrm{L}^2(\mathbb{R}^d)$$

5.
$$\bigcap_{j\in\mathbb{Z}} V_j = \{0\}$$

6. $\{\tau_\lambda\varphi\}_{\lambda\in\Lambda}$ is an ONB for V_0.

At this point, Professor Benedetto provided some motivation for why we need that all eigenvalues of A have norm larger than 1. My notes are particularly weak here, so I am omitting it rather than embarrass myself. I will just move directly to the wavelet system associated with an MRA, as above.

Proposition 14.5 (Y. Meyer, MRA Theorem)**.** *Given an MRA $\varphi, \{V_j\}$ of $L^2(\mathbb{R}^d)$ associated with a dilation matrix and lattice (A, Λ), let $q = |\det(A)|$. Then, there exists $q-1$ constructible functions $\psi_1, \ldots, \psi_{q-1}$ s.t. the wavelet system*

$\{\psi_{\ell,m,\lambda} \mid \ell = 1, \ldots, q-1,\ m \in \mathbb{Z},\ \lambda \in \Lambda\}$ *where* $\psi_{\ell,m,\lambda}(x) = |\det(A)|^{1/2} \cdot \psi_\ell(A^m \cdot x - \lambda)$

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

Comments: To provide a general idea of the direction of the construction, suppose that we have a dilation matrix and lattice, (A, Λ). Let $q = \det(A)$ and suppose that $\{a_\ell\}_{\ell=1,\ldots,q-1}$ is a sequence of functions defined on Λ s.t. the following hold:

1.
$$\sum_{\lambda\in\Lambda} a_j(\lambda) \cdot \overline{a_k(A \cdot \gamma + \lambda)} = q \cdot \delta(j,k) \cdot \delta(\gamma, 0)$$

2.
$$\sum_{\lambda\in\Lambda} a_0(\lambda) = q$$

Then, define
$$\psi_\ell(x) = \sum_{\lambda\in\Lambda} a_\ell(\lambda) \cdot \varphi(A(x) - \lambda),\ \ell = 1, \ldots, q-1.$$

Conditions 1 and 2 on $\{a_\ell(\lambda)\}$ provide the orthonormality of the set of functions $\{\psi_\ell\}$. To prove that $\{\psi_\ell\} \cup \{\psi_0 = \varphi\}$ forms an ONB for V_1 requires significant work. We proceed in much the same way as for the dyadic MRA of $L^2(\mathbb{R})$. We consider the functions

$$\sum_{\lambda\in\Lambda} a_\ell(\lambda) \cdot e^{-2\pi i\lambda\cdot\gamma},\ \gamma \in \mathbb{T}^d \stackrel{\text{def}}{=} \mathbb{R}^d/\Lambda,$$

which turn out to be QMFs.

The remainder of this lecture provides motivation or fills in holes for the above.

Lemma 14.6 (Tiling Lemma)**.** *Let Q be Lebesgue measurable, and assume that*

$$\bigcup_{\lambda\in\mathbb{Z}^d} (Q + \lambda) = \mathbb{R}^d.$$

Then, the following are equivalent:

1. $\forall \lambda \in \mathbb{Z}^d \setminus \{0\}$, $|Q \cap (Q + \lambda)| = 0$.
2. $|Q| = 1$.

Proof. Define the $\mathbb{Z}^d$ periodic function

$$f(x) = \sum_{\lambda \in \mathbb{Z}^d} \mathbb{1}_Q(x - \lambda).$$

Note that

$$|Q| = \int_{\mathbb{R}^d} \mathbb{1}_Q(x)\, dx = \sum_{\lambda \in \mathbb{Z}^d} \int_{[0,1]^d - \lambda} \mathbb{1}_Q(x)\, dx = \sum_{\lambda \in \mathbb{Z}^d} \int_{[0,1]^d} \mathbb{1}_Q(x - \lambda)\, dx = \int_{[0,1]^d} f(x)\, dx$$

By the covering hypothesis, we see that $f(x) \geq 1$ a.e. The two conditions above are clearly seen to both be equivalent to $f = 1$ a.e. □

Proposition 14.7. *If A is a dilation matrix with lattice Λ, then* $|\Lambda / A(\Lambda)| = \det(A)$.

This is proved using the tiling lemma.

Theorem 14.8. *Suppose that we have an MRA $\varphi, \{V_j\}$ for $L^2(\mathbb{R})$ associated with dilation matrix and lattice (A, Λ), and $f \in L^2(\mathbb{R}^d)$. Here, we denote $\mathbb{R}^d / \Lambda$ as $\mathbb{T}^d$. Then,*

$$\frac{1}{|\det(A)|^{1/2}} \cdot f \in V_1 \iff \exists H_f \in L^2(\mathbb{T}^d) \text{ s.t. } |\det(A)|^{1/2} \cdot \hat{f}(A^*(\gamma)) = H_f(\gamma) \cdot \hat{\varphi}(\gamma) \text{ a.e.}$$

In this case, we have $\| H_f \|_{L^2(\mathbb{T}^d)} = \| f \|_{L^2(\mathbb{R}^d)}$.

Proof. We know that

$$\frac{1}{|\det(A)|^{1/2}} \cdot f \in V_1 \iff \frac{1}{|\det(A)|^{1/2}} \cdot f(A^{-1}(x)) \in V_0$$

Then, we know that

$$\frac{1}{|\det(A)|^{1/2}} \cdot f(A^{-1}(x)) = \sum_{\lambda \in \Lambda} h_f[\lambda] \cdot \tau_\lambda \varphi(x) \text{ where } h_f[\lambda] = \langle \frac{1}{|\det(A)|^{1/2}} \cdot f \circ A^{-1}, \tau_\lambda \varphi \rangle$$

Taking the Fourier transform, we get

$$|\det(A)|^{1/2} \cdot \hat{f}(A^*(\gamma)) = \hat{\varphi}(\gamma) \cdot \sum_{\lambda \in \Lambda} h_f[\lambda] \cdot e^{-2\pi i \lambda \cdot \gamma}$$

In the original MRA calculations, we continually split the summation into the even and odd pieces, based on the behavior of the complex exponential. In the above, we can analogously split the summation into the cosets of $\Lambda / A(\Lambda)$ (q pieces). It is much the same, and we are done. □

Now, the last piece in our outlined proof of the MRA theorem.

Proposition 14.9. *Suppose that we have an MRA $\varphi, \{V_j\}$ for $L^2(\mathbb{R})$ associated with dilation matrix and lattice (A, Λ). Let $\psi_0, \dots, \psi_{q-1} \in V_1$. Then,*

1. *The set $\{\tau_\lambda \varphi \mid \lambda \in \Lambda\}$ is ON if and only if*

$$\sum_{\ell=0}^{q-1} |\mathcal{G}_0^r(\gamma)|^2 = q \text{ a.e.}$$

2. *The set $\{\tau_\lambda \psi_\ell \mid \lambda \in \Lambda, \ell = 0, \dots, q-1\}$ is orthonormal if and only if the vectors*

$$V_\ell(\gamma) = \left(\mathcal{G}_\ell^0(\gamma), \dots, \mathcal{G}_\ell^{q-1}(\gamma)\right) \in \mathbb{C}^q$$

   are orthonormal a.e.

3. *The set $\{\tau_\lambda \psi_\ell \mid \lambda \in \Lambda, \ell = 0, \dots, q-1\}$ is an ONB for V_1 if and only if the $q \times q$ matrix $\mathcal{U}(\gamma) = (\mathcal{G}_\ell^r(\gamma))$ is unitary a.e.*

I have no clue what $\mathcal{G}$ is, it snuck by me in the lecture.

References

[1] J.J. Benedetto. *Harmonic Analysis and Applications*. CRC Press, 1996.

[2] J.J. Benedetto and G. Zimmerman. Sampling multipliers and the poisson summation formula. *Journal of Fourier Analysis and Applications*, 3, 1997.

[3] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1992.

[4] I. Daubechies, A. Grossman, and Y. Meyer. Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5):1271–1283, 1986.

[5] P. Enflo. A counterexample to the approximation problem in banach spaces. *Acta Math*, (130):309–317, 1973.

[6] C. Heil. A Basis Theory Primer. 1997.

[7] W.R. Madyck. Some elementary properties of multiresolution analyses of $L^2(\mathbb{R}^n)$. In C. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, pages 259–294. Academic Press, 1992.

[8] S. Mallat. Multiresolution approximation and wavelet orthonormal bases of L2. *Transaction of the American Mathematical Society*, 315:69–87, 1989.

[9] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1999.

[10] Y. Meyer. *Wavelets and Operators*. Cambridge University Press, 1995.

[11] Yves Meyer. Ondelettes, fonctions splines et analyses graduées. *Rend. Sem. Mat. Univ. Politec. Torino*, 45(1):1–42 (1988), 1987.

[12] J. Morlet, A. Grossman, and T. Paul. Transforms associated to square integrable group representations. *Journal of Mathematical Physics*, 26, 1985.

[13] A. Pełczyński. On the impossibility of imbedding the space L in certain banach spaces. *Colloq. Math*, (8):199–203, 1961.

[14] W. Rudin. *Real & Complex Analysis*. McGraw-Hill Series in Higher Mathematics, third edition, 1987.

[15] P.P. Vaidyanathan. *Multirate systems and filter banks*. Prentice Hall, 1993.

[16] R.M. Young. *An Introduction to Nonharmonic Fourier Series*. Academic Press, 1980.