# Time-Frequency Analysis and the Wavelet Transform

Ryan Plestid

August 19th, 2013

# Overview of Talk

- Non periodic signals
- The Windowed Fourier Transform
- The Continuous Wavelet Transform
- The Discrete Wavelet Transform
- The Wigner Distribution Function
- Implementing Algorithms on the Computer
- Using proprietary software for data analysis

# Non Stationary Signals

- Often times a signal will not have a time independent frequency. Take for example a short burst of EM radiation. After some finite interval of time the signal would vanish.
- To determine when a signal occurs as well as what normal modes are most prevalent in its construction requires time-frequency or time-scale analysis
- There are two major classes of transforms for this Short Time Fourier Transform and the Wavelet Transformation

# Solution One the Windowed Fourier Transform

$$\boxed{\tilde{f}(\omega, \tau) = \int e^{2\pi i \omega t} \, \overline{w}(t-\tau) \, f(t) \, \mathrm{d}\tau} \qquad (1)$$

- Exactly like the normal Fourier transform except that the signal and the complex exponential are convolved with a function with compact support.
- This solves the problem of temporal resolution however one must remember that due to the uncertainty principle associated with Fourier Analysis there is an intrinsic "trade off" between frequency and time resolution.

# Proof of the Fourier Uncertainty Principle

Assume that $f$ is a normalizable function i.e. $f \in L^2$

$$\int f^2(x)\mathrm{d}x \in \mathbb{R}$$

then by integrating by parts, and noting that $x \cdot |f|^2$ goes to zero at positive and negative infinity we can say that

$$|f|^2_{L^2} = \int_{\mathbb{R}} |f|^2 = -\int_{\mathbb{R}} x(f \cdot \overline{f})' = -2Re\int_{\mathbb{R}} x \cdot f \cdot \overline{f}\ '$$

Next we use that a number's modulus will always be greater than or equal to it's real part's modulus as well as employing the Cauchy Schwartz inequality.

$$\left|-2Re\int_{\mathbb{R}} x \cdot f \cdot \overline{f}\ '\right| \ \leq \ 2\int_{\mathbb{R}} \left|x \cdot f \cdot \overline{f}\ '\right| \ \leq \ 2 \cdot |xf|_{L^2} \cdot |f'|_{L^2}$$

# Proof of the Uncertainty Principle (contd)

Now if we use Parsavel's Identity and note that the derivative of a Fourier transform is equivalent to multiplication by $2\pi\xi$

$$|f'|_{L^2} = \left| \hat{f}\,' \right|_{L^2} = \left| 2\pi\xi \cdot \hat{f} \right|_{L^2}$$

So we now have that

$$|f|_{L^2}^2 \leq 4\pi \cdot |xf|_{L^2} \cdot |\xi \hat{f}|_{L^2}$$

Next we need to turn this result into a statement involving variance instead of just the second moment. Define $g \equiv e^{-2\pi i x \xi_0} f(x + x_0)$ then $\hat{g} = e^{2\pi i x_0 \xi} f(\xi + \xi_0)$. Applying the inequality above to $g$ and $\hat{g}$ gives

$$|f|_{L^2}^2 \leq 4\pi \cdot |(x - x_0) f|_{L^2} \cdot |(\xi - \xi_0)\hat{f}|_{L^2}$$

# Minimizing the Uncertainty Principle

- To minimize product of the uncertainties in both the real and reciprocal space one can use a Gaussian wave packet.
- The Fourier transform of a Gaussian wave packet is another Gaussian wave packet and the product of their standard deviations results in a saturation of the uncertainty principle.
- For this reason the window function is almost always chosen to be a Gaussian to help optimize time-frequency resolution.

# Reconstructing the Signal

- In Fourier analysis we know that there is a one to one correspondence between a signal and its Fourier transform. But does such an inverse function exist for the windowed Fourier transform.
- This inverse function may not necessarily be as symmetric as the Fourier transform's inverse and it will also map from a function of two variables (time and frequency) to a function of one variable.
- To derive this reconstruction formula we must use the Fourier transform to our advantage.

# Reconstruction Formula

We begin by noting that the windowed Fourier transform is just the regular Fourier transform of our signal multiplied by a window function

$$\tilde{f}(\omega, \tau) = \mathscr{F}\{f(t)\overline{w}(t-\tau)\}$$

Where we have defined $f_w$ as product of the signal and window function. Next we use the inverse Fourier transform

$$\overline{w}(t-\tau) = \mathscr{F}^{-1}\left\{\tilde{f}(\omega, \tau)\right\}$$

Here you might expect we are done however we cannot simply divide by the window function because it might vanish. If we can generate the norm of the window function we can divide by it because it must have a non zero norm

# The Reconstruction formula cont'd

So next Then multiplying both sides by the window function and integrating with respect to tau we obtain

$$f(t)\int_{\mathbb{R}} \mathrm{d}\tau \left|w(t-\tau)\right|^2 = \int_{\mathbb{R}} \mathrm{d}\tau w(t-\tau)\mathscr{F}^{-1}\left\{\tilde{f}(\omega,\tau)\right\}$$

$$\boxed{f(t) = \frac{1}{||w||_{L^2}^2}\iint_{\mathbb{R}^2} \mathrm{d}\omega \mathrm{d}\tau \; e^{2\pi i\omega t} w(t-\tau)\tilde{f}(\omega,\tau)} \qquad (2)$$

This is the reconstruction formula for the windowed Fourier transform. It is relatively simple and has a straight forward derivation however this "practice proof" will help in understanding how to obtain the wavelet transform's reconstruction formula.

# Review of the Windowed Fourier Transform

- The WFT can resolve both oscillatory and temporal details
- Resolutions satisfy the Heisenberg uncertainty principle (minimized by Gaussian windows)
- Somewhat computationally expensive
- Uses the same sized "filter window" for all frequencies

# Multi-Resolution Analysis

- One could imagine that at high frequencies one would prefer better temporal resolution while at low frequencies improved oscillatory resolution would be desired.
- The WFT fails to achieve this. You must pick a window and stick to it throughout the transform. This is undesirable
- This determines the time-frequency resolution trade off for the entire duration of the transform
- If the window function naturally scaled it's self to provide desirable oscillatory and temporal resolution for whichever frequency was of current interest them **multi-resolution analysis** could be executed.

# The Continuous Wavelet Transformation

$$\boxed{\tilde{f}(s,t) = \int_{\mathbb{R}} \mathrm{du}\ f(\mathrm{u})\overline{\psi}\left(\tfrac{\mathrm{u}-t}{s}\right)|s|^{-p}} \qquad (3)$$

The continuous wavelet transform is one example of the aforementioned multi-resolution analysis, here $s$ is scale and $t$ is time. The wavelet $\psi$ must satisfy the following conditions.

- The wavelet must have compact support (be non zero exclusively on some bounded interval)
- $0 < \int_{\mathbb{R}} \frac{\left|\hat{\psi}(x)\right|^2}{|x|} \mathrm{d}x < \infty$ (this will be justified later)
- $\int_{\mathbb{R}} \psi(\mathrm{u})\mathrm{du} = 0$ (This follows from above)

NOTE: Here $p$ is left general, the popular choice is $p = 1/2$ this is important if you read other literature.

# Concept of the Mother Wavelet

- The wavelet $\psi$ is assumed to be a function of just $u$ and this is called the mother wavelet as it generates all of the daughter wavelets
- The distinction between the two is that the mother wavelet is static while the daughter wavelets are **scaled** and **translated in time**
- This can also be explained as $\psi = \psi(\mathrm{u})$ and $\psi_{s,t} = \psi(\frac{\mathrm{u}-t}{s})|s|^{-p}$ where the former is the mother wavelet, and the latter is a scaled and translated wavelet
- Note that $\psi_{1,0} = \psi$

# Deriving the Reconstruction Formula Part 1

We begin by treating all of our functions as vectors in $L^2$ space for brevity's sake also note that here the inner product is the standard integral formulation for functional spaces. Employing Parsavel's identity we note that

$$\tilde{f}(s,t) = \langle \psi_{s,t} | f \rangle = \left\langle \hat{\psi}_{s,t} \middle| \hat{f} \right\rangle$$

$$\hat{\psi}_{s,t}(\omega) = \int_{\mathbb{R}} \mathrm{du}\ |s|^{-p} \psi\left(\tfrac{\mathrm{u}-t}{s}\right) e^{-2\pi i \omega \mathrm{u}}$$

Note that here u is time and $\omega$ is frequency now defining $\mathrm{u}' = \frac{\mathrm{u}-t}{s}$ we can note that the above is equivalent to

$$\int_{\mathbb{R}} \mathrm{du}' |s|^{1-p} \psi(\mathrm{u}') e^{-2\pi i \omega s(\mathrm{u}'+t)} = |s|^{1-p} e^{-2\pi i \omega t} \hat{\psi}(s\omega)$$

# Deriving the Reconstruction Formula Part 2

Using this result and taking the inner product with $\hat{f}$ we obtain

$$\left\langle \hat{\psi}_{s,t} \middle| \hat{f} \right\rangle = |s|^{1-p} \int_{\mathbb{R}} \mathrm{d}\omega e^{2\pi i \omega t} \overline{\hat{\psi}}(s\omega)\hat{f}(\omega)$$

$$= |s|^{1-p} \mathscr{F}_t^{-1} \left\{ \overline{\hat{\psi}}(s\omega)\hat{f}(\omega) \right\}$$

Note the subscript on the inverse Fourier transform; this denotes that here the time variable was $t$ not u.ext we equate what is above to $\hat{f}$ and take the Fourier transform (with respect to $t$ **not** u of both sides.

$$\mathscr{F}_t \left\{ \tilde{f}(s,t) \right\} = |s|^{1-p} \overline{\hat{\psi}}(s\omega)\hat{f}(\omega)$$

# Deriving the reconstruction formula Part 3

Note we cannot divide by $\overline{\hat{\psi}}(s\omega)$ because it might vanish, however if an expression can be obtained for $\hat{f}$, $f$ can be obtained trivially. We start by multiplying both sides of the equation by $w(s)$ and $\hat{\psi}(s\omega)$then integrating over all positive values of $s$

$$\int_{\mathbb{R}^+} \mathrm{d}s\ w(s) \int_{\mathbb{R}} \mathrm{d}t\ e^{-2\pi i t\omega} \hat{\psi}(s\omega)\tilde{f}(s,t) =$$

$$\int_{\mathbb{R}^+} \mathrm{d}s\ w(s)|s|^{1-p}\hat{f}(\omega)\left|\hat{\psi}(s\omega)\right|^2$$

Now we define $Y(\omega) \equiv \int_{\mathbb{R}} \mathrm{d}s\ w(s)|s|^{1-p}\left|\hat{\psi}(s\omega)\right|$ we will assume this function is positive, bounded, and non-zero almost everywhere (a.e. for short); this ensures we can divide by $Y(\omega)$.

# Deriving the reconstruction formula Part 4

Now we begin with another definition $\hat{\psi}^{s,t} \equiv Y(\omega)^{-1}\hat{\psi}_{s,t}(\omega)$ $\{\psi^{s,t}\}$ are called reciprocal wavelets. Taking advantage of these new definitions we can write

$$\hat{f}(\omega) = Y(\omega)^{-1} \int_{\mathbb{R}^+} \mathrm{d}s \; w(s) \int_{\mathbb{R}} \mathrm{d}t \; e^{-2\pi i \omega t} \hat{\psi}(s\omega) \tilde{f}(s,t)$$

$$= Y(\omega)^{-1} \int_{\mathbb{R}^+} \mathrm{d}s \; w(s) s^{p-1} \int_{\mathbb{R}} \mathrm{d}t \; \hat{\psi}_{s,t}(\omega) \tilde{f}(s,t)$$

$$= \int_{\mathbb{R}^+} \mathrm{d}s \; w(s) s^{p-1} \int_{\mathbb{R}} \mathrm{d}t \; \hat{\psi}^{s,t}(\omega) \tilde{f}(s,t)$$

Now taking the inverse Fourier transform of both side we obtain

$$\boxed{f(\mathrm{u}) = \int_{\mathbb{R}^+} \mathrm{d}s \; w(s) s^{p-1} \int_{\mathbb{R}} \mathrm{d}t \; \psi^{s,t}(\omega) \tilde{f}(s,t)} \qquad (4)$$

# The Reciprocal Wavelets

In the previous slide we defined this new set of reciprocal wavelets for our convenience. Here we will investigate some of their properties because of their inclusion in the reconstruction formula.

$$\psi^{s,t}(\mathrm{u}) = s^{1-p} \int_{\mathbb{R}} \mathrm{d}\omega \; e^{2\pi i \omega(\mathrm{u}-t)} Y(\omega)^{-1} \hat{\psi}(s\omega) = \psi^{s}(\mathrm{u}-t)$$

Next we will try and see how scaling is affected

$$\psi^{s}(\mathrm{u}) = s^{1-p} \int_{\mathbb{R}} \mathrm{d}\omega \; e^{2\pi i \omega \mathrm{u}} Y(\omega)^{-1} \hat{\psi}(s\omega)$$

$$= s^{-p} \int_{\mathbb{R}} \mathrm{d}\omega \; e^{2\pi i \omega \mathrm{u}/s} Y(\tfrac{\omega}{s})^{-1} \hat{\psi}(\omega)$$

So it becomes clear that *iff* $Y(\omega) = Y(\frac{\omega}{s})$ a.e. then $\{\psi^{s,t,}\}$ can be generated from the mother wavelet $\psi^{1}$

# Scale invariance

Remember that when defining $Y(\omega)$ we included the yet undetermined $w(s)$. Now we realize our goal is to make $Y(\omega) = Y(\frac{\omega}{s})$ a.e. Let $w(s) = s^{p-2}$ then

$$\mathrm{d}s \ w(s)s^{1-p} = \frac{\mathrm{d}s}{s}$$

Now making the substitution $s\omega = \pm z$

$$Y(\omega) = \int_{\mathbb{R}^+} \frac{\mathrm{d}s}{s} \left|\hat{\psi}(s\omega)\right|^2 = \int_{\mathbb{R}^+} \frac{\mathrm{d}z}{z} \left|\hat{\psi}(\pm z)\right|^2 \equiv C_\pm$$

Note that this holds everywhere except at 0 (a set of measure 0) and that $Y(\omega)$ is now piecewise constant and therefore $Y(\omega) = Y(s\omega)$ and therefore $\{\psi^{s,t}\}$ can be generated from $\psi^{1,0}$. We also impose the condition that $0 < C_\pm < \infty$ this is known as the admissibility condition.

# A more intuitive admissibility condition

If $\psi$ is an admissible wavelet then it satisfies

$$\int_{\mathbb{R}^+} \frac{\mathrm{d}z}{z} \left|\hat{\psi}(\pm z)\right|^2 < \infty$$

Note that unless as $t \to 0$ $\hat{\psi}(t) \to 0$ then the integral will not be finite. So assuming $\hat{\psi}$ is continuous then $\hat{\psi}(0) = 0$ this is equivalent to saying

$$\int_{\mathbb{R}} \psi(t) e^{2\pi i t \cdot 0} \mathrm{d}t = \int_{\mathbb{R}} \psi(t) \mathrm{d}t = 0$$

That is to say that a wavelet must be a wave centered about the $t$ axis.

# A special case: Statistical Distributions

Suppose we have a function $p(u)$ such that

$$\int_{-\infty}^{\infty} p(u)\mathrm{d}u = 1 \quad \int_{-\infty}^{\infty} u\ p(u)\mathrm{d}u = 0 \quad \int_{-\infty}^{\infty} u^2 p(u)\mathrm{d}u = 1$$

(i.e. a probability distribution with 0 mean and unit variance) then let us assume that $p(u)$ is $n$ times differentiable where $n \geq 1$. Then assume that

$$\lim_{u\to\pm\infty} p^{(n-1)}(t) = 0$$

Then let

$$\psi^n(u) \equiv (-1)^n p^{(n)}(u)$$

And assume that $\hat{\psi}^n$ is continuous then we have that

$$\int_{-\infty}^{\infty} \mathrm{d}t\psi^n(u) = (-1)^n \left(p^{(n-1)}(\infty) - p^{(n-1)}(-\infty)\right) = 0$$

# Statistical Distribution part 2

And thus $\psi^n$ is an admissible mother wavelet not let us define

$$\psi^n_{s,t} \equiv s|^{-1}\psi^n\left(\frac{u-t}{s}\right) \quad p_{s,t}(u) \equiv |s|^{-1}p\left(\frac{u-t}{s}\right)$$

Here $p_{s,t}$ is a statistical distribution with $\langle u \rangle = t$ and $\sigma = s$

$$\overline{f}(s,t) \equiv \langle p_{s,t}|f\rangle \quad \tilde{f}_n(s,t) \equiv \langle \psi^n_{s,t}|f\rangle$$

Here $\overline{f}$ is a local average of the signal. Now noting that

$$\psi^n_{s,t} = (-1)^n s^n \partial^n_u p_{s,t}(u) = s^n \partial^n_t p_{s,t}(u)$$

We have that

$$\tilde{f}_n(s,t) = \int_{\mathbb{R}} \mathrm{d}u\ \psi^n_{s,t} = s^n\partial^n_t \int_{\mathbb{R}} \mathrm{d}u\ p_{s,t}(u)f(u) = s^n\partial^n_t\overline{f}(s,t)$$

i.e the CWT is proportional to the average's n$^{th}$ derivative

# Examples of wavelets

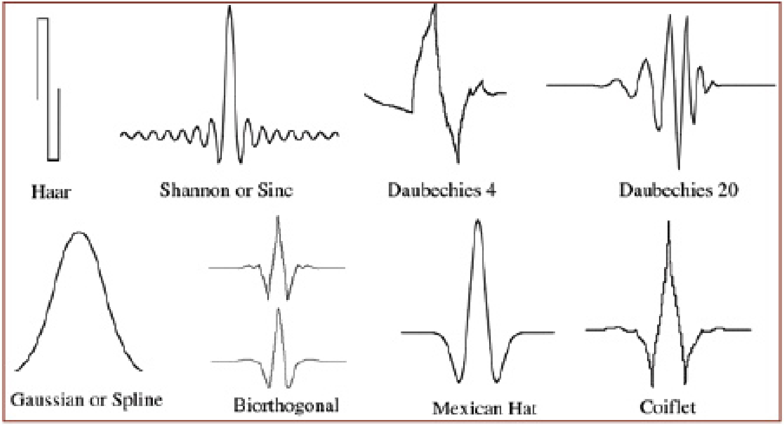



*Figure 8*

*Examples of types of wavelets*

Note: The Mexican hat wavelet is the second derivative of a Gaussian statistical distribution

# Concepts of the Continuous Wavelet Transform

- The wavelet scales to modify its temporal and frequency resolution to better resolve the signal
- Wavelets must satisfy particular properties
- The wavelet is naturally expressed using scale, not frequency and the two are related by inversion
- Enough information is preserved by the CWT to ensure that the original signal can be recovered. It could conceivably (and in fact does) contain redundancies

# Discrete Wavelet Transform

$$\mathbf{f} = \sum_{k\in\mathbb{Z}}\sum_{j\in\mathbb{Z}} c_{j,k}\ \mathbf{w}_{j,k} \quad c_{j,k} = \langle \mathbf{f}\,|\mathbf{w}_{j,k}\rangle \tag{5}$$

- Vector notation has been used here to stress this is a functional space. So $\mathbf{w}_{j,k} \equiv w_{j,k}(t) = 2^{j/2} w(2^j t - k)$ where $\{\mathbf{w}_{j,k}\}$ are the wavelet notes.
- This transform is analogous to the Fourier Series when it is compared to the Fourier Transform.
- The notes actually forms a basis (i.e. they are linearly independent and span the space)
- By creating a Basis we obtain the minimum amount of information required to reconstruct the signal

# Discretized Multi-Resolution Analysis Part 1

Let us examine the Hilbert Space of $L^2$ with the standard functional inner product (i.e. Lebesque measure). $\{\mathbf{w}_{j,k}\}$ forms a basis and therefore we can write any function $\mathbf{f} \in L^2$ as

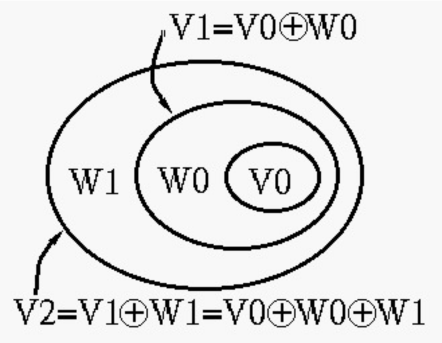
$$\mathbf{f} = \sum_{k \in \mathbb{Z}} \sum_{j \in \mathbb{Z}} c_{j,k} \; \mathbf{w}_{j,k}$$

Now let us partition the set of all functions into cells $W_i$ such that $W_N$ is the space which is spanned by the set of linearly independent functions $\{\mathbf{w}_{N,k}\}$ with $N$ fixed and $k \in \mathbb{Z}$. Similarly we can define $V_N$ to be the space spanned by $\{\mathbf{w}_{i,k}\}$ such that $i < N$ and $k \in \mathbb{Z}$.

# Discretized Multi-Resolution Analysis Part 2

It is relatively straight forward to see that $V_N = V_{N-1} \oplus W_N$ because $W_N$ contains all of the new linearly independent vectors that appear in $V_N$ but were not in $V_{N-1}$ also the fact that $W_N \neq \emptyset \;\; \forall N \in \mathbb{Z}$ this implies that

$$\emptyset \subset \ldots V_{N-1} \subset V_N \subset V_{N+1} \ldots \subset L^2$$



# Discretized Multi-Resolution Analysis Part 3

Each of the nested subspaces is composed from functions of a particular scale (and frequency). Now let $\phi \in V_N$ then

$$\phi = \sum_{k\in\mathbb{Z}}\sum_{j<N} c_{j,k}\ \mathbf{w}_{j,k} = \underbrace{\sum_{k\in\mathbb{Z}}\sum_{j<N-1} c_{j,k}\ \mathbf{w}_{j,k}}_{a_1\in V_{N-1}} + \underbrace{\sum_{k\in\mathbb{Z}} c_{j,k}\ \mathbf{w}_{N,k}}_{d_1\in W_N}$$

Now we have that $\phi = a + d_1$ if continued deconstructing the sum above we could eventually end up with

$$\phi = a_\alpha + \sum_{i=1}^{\alpha} d_i$$

$a$ is known as the analysis or average coefficient and $\{d\}$ are the detail coefficients.

# Admissible wavelet notes

- Notes must be orthogonal i.e. $\langle \mathbf{w}_{j,k} | \mathbf{w}_{j',k'} \rangle = \delta_{j,j'} \delta_{k,k'} C_{j,k}$
- The set of all notes must form a basis for $L^2$ i.e
  $\forall \ \mathbf{f} \in L^2 \ \ \exists ! \{c_{j,k}\} \ \ s.t. \ \ \mathbf{f} = \sum_{-\infty}^{\infty} \sum_{-\infty}^{\infty} c_{j,k} \mathbf{w}_{j,k}$
- $\mathbf{w}_{j,k} \in L^2$ and $\int_{\mathbb{R}} \mathrm{d}t \ w_{j,k}(t) = 0 \ \ \forall j, k \in \mathbb{Z}$

One can construct all of $L^2$ using the wavelet functions $\mathbf{w}_{j,k}$ provided you allow any possible integer value for $j$ or $k$. This has the distinct disadvantage of requiring an infinite number of coefficients which is impossible to implement numerically.

# Scaling Functions

Let us turn back to the nested subspaces discussed previously and attempt to construct something that can serve to amend this inconvenience. Let us construct an alternate set of functions which form a basis for $L^2$ We will require that

- $\{\phi_{j,k}\}$ forms a basis for $L^2$ (see previous slide for definition) where $\phi_{j,k} = 2^{j/2}\phi(2^j t - k)$
- $\phi_{J,k}$ for a fixed $J$ is orthogonal to its integer translations i.e. $\langle \phi_{J,k} | \phi_{J,l} \rangle = \delta_{k,l} C_j$
- $\phi_{j,k} = \sum_{\ell} h_\ell \; \phi_{j+1,\ell} \;\; \forall j,k \in \mathbb{Z}$

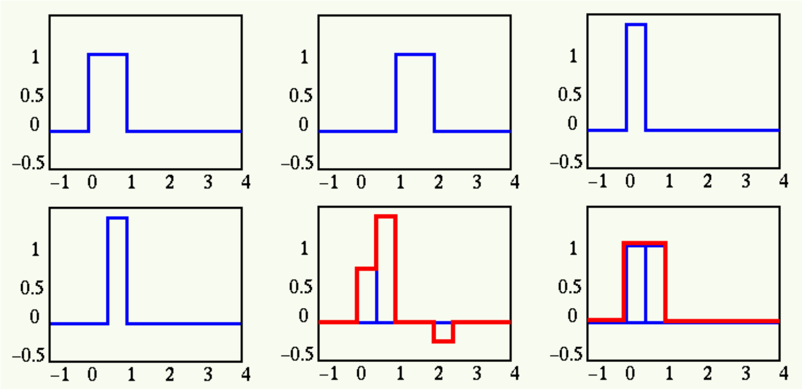
# Combining the Scaling and Wavelet families

The final condition is the most important, it states that any scaling function at a resolution $j$ can be written as a sum of the resolution above it. This is a cascading effect so that $\{\phi_{J,k}\}\ \ \forall k \in \mathbb{Z}$ form a basis for $V_J$. So $\phi_J$ acts as a plug $\forall\ W_j\ \ s.t.\ \ j < J$ so that we need not wander down to $V_{-\infty}$

During the analysis of a discretely sampled signal one is restricted to the range of frequencies from 0 to the Nyquist frequency. This upper bound removes the need to wander up to $V_\infty$. As a result with the combined family of wavelet functions, and scaling functions one can obtain a complete description of a signal.

The more wavelet coefficients that are used the more "detail" or high frequency oscillations one recovers from a signal

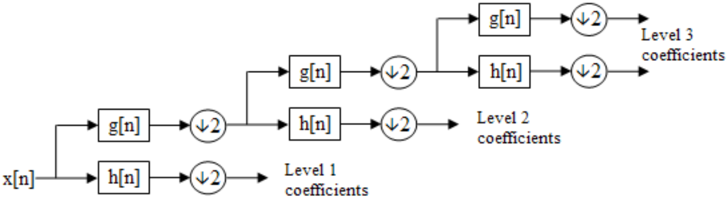
# An example of a scaling function




Here are the Haar scaling functions at two different scales. Note these are not admissible wavelets but as a result they can generate any Haar wavelets of greater or equal scale. The final image shows integer translations of $\phi_{j,0}$ being used to generate $\phi_{j+1,0}$.

# Hidden filters And Filter Banks

If one thinks for a bit about what subspaces are spanned by the scaling functions and wavelet functions it becomes relatively obvious that they are low-pass and band-pass filters respectively when convoluted with a signal. This allows them to be used as a filter bank.




Here g[n] are the analysis coefficients, and h[n] are the detail coefficients. Essentially the signal is passed through a high frequency band pass filter and separated into analysis and detail coefficients. The process is repeated iteratively on the analysis coefficient until enough detail has been removed.

# The efficiency of the discrete wavelet transform

Computationally the DWT uses the least amount of information possible required to reconstruct the signal. After each level in the decomposition it halves the sampling rate because the new Nyquist frequency is half the old one. This is known as **dyadic sampling**.

This gives the DWT the following features

- Can be computed in $N$ time compared to the FFT which is calculated in $Nlog(N)$ time
- Uses the minimum amount of information (good for data compression bad for spectrographs)
- Significantly easier to implement numerically than CWT

# Numerical Implementation of the Continuous Wavelet Transform

Obviously nothing can be implemented continuously on a computer and so the major difference between the CWT and the DWT is the rate of sampling throughout the transform. For the CWT no dyadic sampling is used and inefficiency is chosen to allow for improved quality of data presentation.

This begs the question: which one is better for which scenarios? In my opinion for data analysis the continuous wavelet transform is superior however for image compression, denoising and extremely quick computations the discreet transform is better.

# Software Platforms

For the Discrete Wavelet Transform

- Jwave
- Pywavelets
- MATLAB

For the Continuous Wavelet Transform

- MATLAB

I attempted to write my own continuous wavelet transform, it provided decent results however MATLAB's transform executed faster and provided better plots

# Using Matlab for Wavelet Analysis

- Matlab has a number of functions in it's wavelet toolbox
- These include the discrete wavelet transform, the wavelet packet transform, and the continuous wavelet transform
- For our purposes the continuous wavelet transform (CWT) is best suited

# Some MATLAB basics Part 1

- MATLAB is designed to work with matrices and vectors. It has highly optimized operations which are designed to work with vectors and matrices. All of these are preceded by a period ".". This just means element by element operation.
- For example if A and B have the same dimensions C=A.*B would multiply the ith element of A by the ith element of B and return a vector C of the same length as A & B.
- All functions in MATLAB can take vectors and return vectors.

# Some MATLAB basics Part 2

- To create a linearly spaced vector you type X:Y:Z where X is the first value in the vector,Y is the spacing, and Z is the last value
- You can also stitch vectors together by using A=[B,C] where B and C are vectors this would make A be composed of all the elements of B followed by all the elements of C.

# Some MATLAB functions

- plot(x,y) given two vectors of the same length (x and y) this will plot the values of one against the other and output an image of said plot
- One can also change the title, and axis labels etc. by including commands after the plot command. (i.e. title('X vs Y'))
- To read data from a text file use textread the syntax is as follows

  [vec1,vec2,...]=textread('filename', FORMAT)

# Some more MATLAB functions

- To read data from a text file which is formatted one can also use textscan the syntax is as follows
  Q=textread(fileID, FORMAT)
  Here Q is a cell matric which means that it

# Basic Wavelet Syntax

- The function *cwt* takes 3 basic arguments, a signal vector, a scales vector, and a wavelet identifier string. It then outputs a two dimensional matrix which can be interpreted as a scaleogram
- Basic syntax is as follows

  $$\text{A} = \text{cwt}(\textit{signal}, \textit{scales}, \textit{'wavname'})$$

- Sometimes frequency is desired instead of scales. to create a freqeuncy vector that corresponds to the scales vector used you can use the function *scal2frq* the syntax is

  $$\text{freq} = \text{scal2frq}(\textit{scales}, \textit{'wavname'}, \textit{sampling rate})$$

# Plotting options and syntax

- For the most flexible plotting options I suggest using the funciton *contourf* this generates a filled contour plot of a two dimensional matrix. You can also include axis vectors the number of contour levels you would like.
- The syntax for this is pretty basic

  $$\text{contourf}(xaxis, \ yaxis, \ matrix, nLVL)$$
- Axis labels and titles can be added in the same way as with the plot command by adding lines to a matlab script after the plot command was called

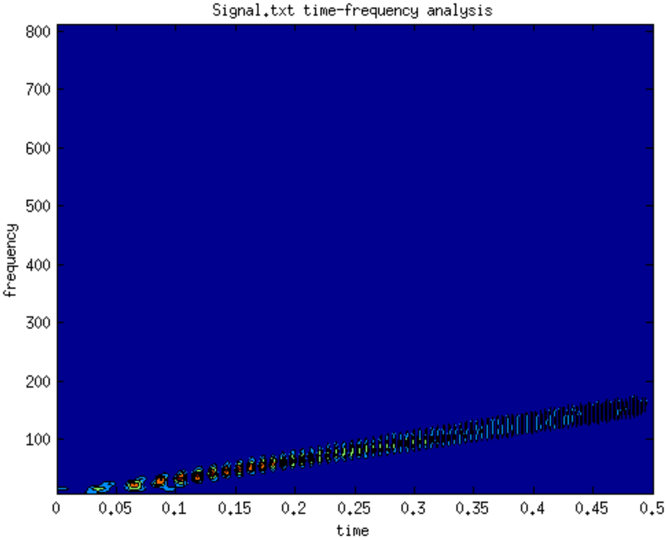
# An example script

```
y=0:0.001:1;
x=sin(y .* y * 600 * pi);
scales=1:100;
scales=1./scales;
scales=wrev(scales); // this just flips the vector
scales=[scales,2:1:100];
A=cwt(x,scales,'morl');
B=A .* A;
freq=scal2frq(scales,'morl',0.001);
contourf(y,freq,B);
title('Linear Chirp frequency and time');
xlabel('Time');
ylabel('Frequency');
```

# Another example script

```
[y,x]=textread('signal.txt','%f%f'); scales=1:100;
 A=cwt(x,scales,'morl');
B=A .* A;
freq=scal2frq(scales,'morl',0.001);
contourf(y,freq,B,30);
title('Signal.txt's time frequency analysis');
xlabel('Time');
ylabel('Frequency');
```

# Output of previous Program



# Notes on Data Analysis

- **contourf** uses relative heights for colouring so if you have one very dominant signal it will wash out smaller ones and assign them the same colour as 0.
- To fix this one can either specify a large number of levels, or analyze smaller sections of the sample individually
- Often if one is not careful about the scale vector used important features can be missed. It is worth using a large scale range (1,100) and if the edges clearly contain no more features then one can zoom in by adjusting the scales that are used i.e. (1,20)
- If the wavelet transform is not squared there are often residual "ripples". I suggest squaring the transform

# Appendix 1: Measure Theory and Lebesgue Integration

A measure is just a generalization of the concept of length,area,volume and hyper volume. It is defined for some set $X$ as any function $\mu : X \to \mathbb{R}$ such that

- For a finite collection of pairwise disjoint sets $\mu(\cup_{i\in\mathbb{Z}} Y_i) = \sum_{i\in\mathbb{Z}} \mu(Y_i)$
- $\mu(\emptyset) = 0$
- $\mu(X) \geq 0 \quad \forall X$

The basic idea of Lebesgue integration is that you take the target set of a function and decompose it into subsets $X_i$ and then find the measure of $I(X_i)$ and sum those together. This allows one to integrate discontinuous functions. The canonical example is a function which is 1 for rational numbers and 0 for irrational numbers.

# Apendix 2: The Wigner Distribution Function

$$\boxed{W(f,t) = \int_{\mathbb{R}} x(t+\tau/2)\cdot\overline{x}(x-\tau/2)e^{-2\pi i f\tau}\,\mathrm{d}\tau} \qquad (6)$$

This transform is interesting in that it provides fantastic spatial and time resolution and is somewhat easier to work out analytically for example let $x(t) = e^{2\pi i f \;\; kt^2}$ then

$$W = \int e^{2\pi i f k(t+\tau/2)^2} e^{-2\pi i f k(t-\tau/2)^2} e^{-2\pi i f\tau}\,\mathrm{d}\tau$$

$$= \int e^{4\pi k t\tau} e^{-2\pi i\tau f}\mathrm{d}\tau = \delta(f-2kt)$$

This would seem to break the uncertainty principle however if one carefully examines a chirp signal one will realize it does not satisfy the conditions necessary in the proof. Namely it is not square integrable.

The more recent Gabor-Wigner transform is worth examining

# references

- Friendly guide to wavelets G. Kaiser
- http://fourier.eng.hmc.edu/e161/lectures/wavelets/node3.html
- 