

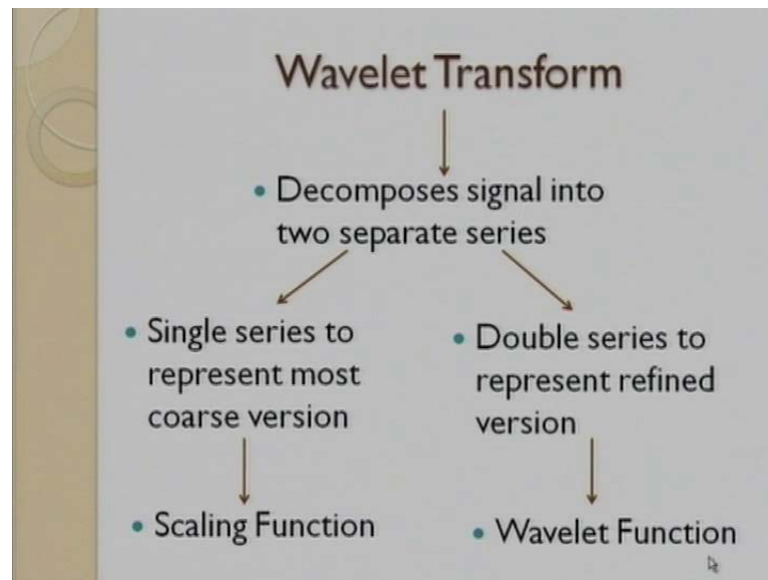
Advanced Digital Signal Processing - Wavelets and Multirate
Prof. V. M. Gadre
Department of Electrical Engineering
Indian Institute of Technology, Bombay

Lecture No # 48
Towards selecting wavelets through vanishing moments

Hello and welcome to this next lecture, in the series on the subject of wavelets and multirate DSP. And in broader sense, the subject that deals with joint time frequency analysis. Today's lecture is titled as, towards selecting wavelets through vanishing moments. So, in this particular lecture, we are going to understand the significance of vanishing moments. However, before we start our journey, let us spend few minutes in recalling what we did last time, because after all what we are going to do in this particular lecture, it in a way stands on the shoulder of what we did in the last couple of lectures.

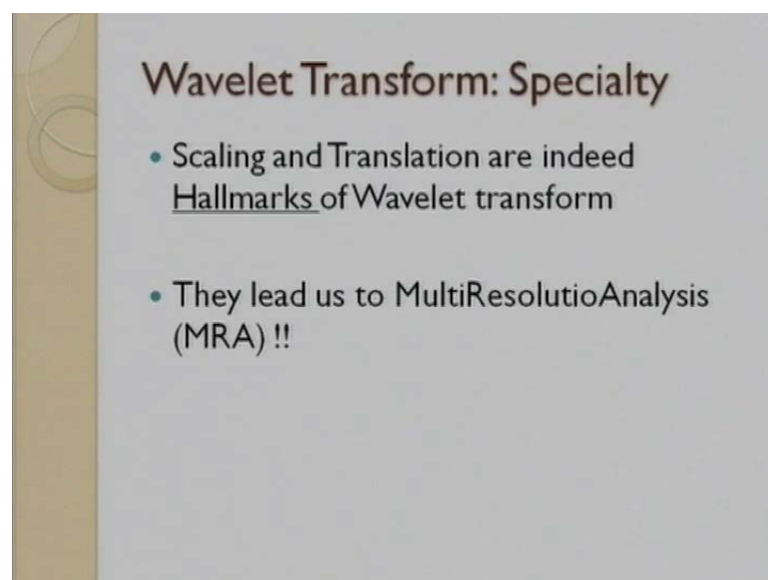
So, let us quickly revisit, what we did in last couple of lectures. If you remember in the last to last lecture we mentioned, and we in a way expressed our belief, that if last hundred years, where the hundred years of Fourier transform. Then probably next hundred years are going to be the hundred years of wavelet transform. And this belief comes out of the fact, that when it comes to dealing with non-stationary type of signals, Fourier transform has its own limitations, and that is where, we will have to go beyond the purview of Fourier transform, and start looking into the characteristics of wavelet transform.

(Refer Slide Time: 02:14)



We started our journey with a basic fundamental principle; that wavelet transform can decompose an underlying signal into two separate series; a single series to represent the most coarse version of the information, and that leads us to scaling function, and double series to represent the refined version, or that gives us the details, which are present in the underlying signal or function, and that leads us to wavelet function.

(Refer Slide Time: 02:49)



We furthered on this, and then we realized; that when it comes to wavelet transform, scaling, translation and dilation, they are indeed the hallmarks of wavelet transform, and

they together need us to MultiResolution analysis, which is very popularly also known as MRA.

(Refer Slide Time: 03:11)

Relationship

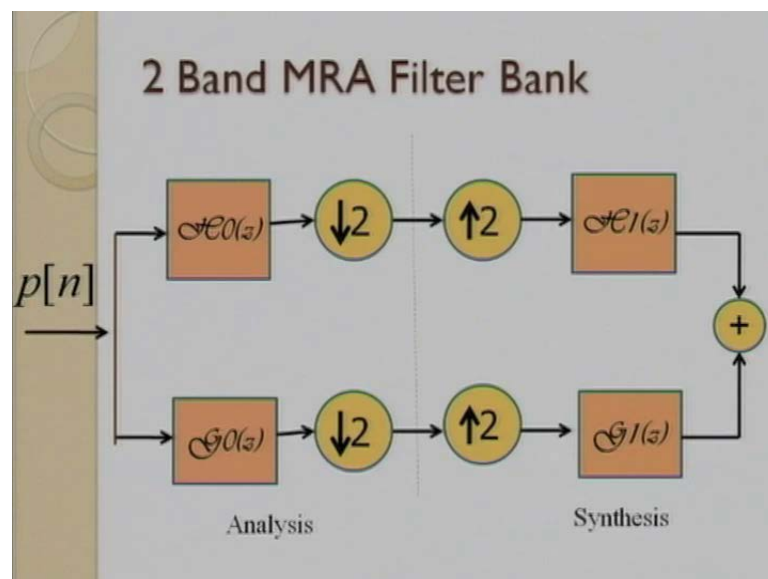
- As the spaces and spans are clear now
- Intuitively, we observe a relationship between these spaces!

$$\dots\dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \dots\dots$$

- Intuitively we can see that as we move towards right, i.e. **up the ladder**, we are moving towards $L_2(\mathbb{R})$

The whole philosophy of MultiResolution Analysis is in a way based on the underlying framework. And this framework comes out of the presence of nested subsets, and these nested subsets, they in a way give us the entire MultiResolution framework.

(Refer Slide Time: 03:39)

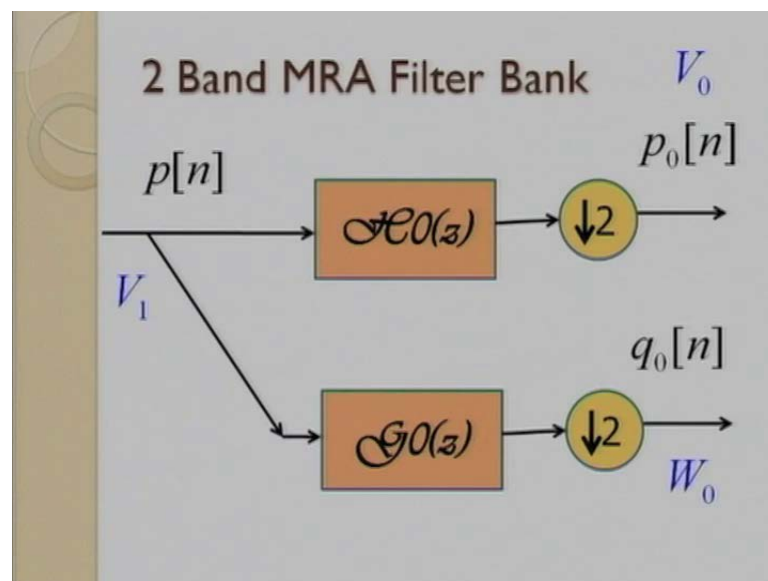


We also looked at, the two band MRA filter bank, and this structure is very typical. For any input signal let us say p of n , if in the analysis side, I have analysis low pass filter,

presented by h_0 of z , and analysis high pass filter, given by g_0 of z , and they are followed by down samplers or disseminators. This is quite essential to keep total number of samples same. For example, if I am starting off with four samples, then I should be able to produce two samples here, and two samples here.

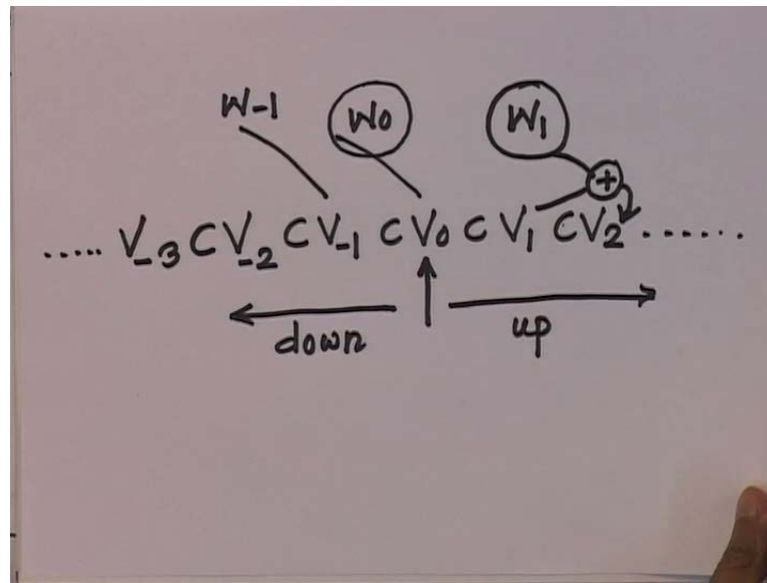
However, the moment you introduce disseminator, or down sampler, the whole structure becomes something that can be called as a multirate structure. The sampling rate is not uniform, and that is why for all the unirate DSP structures, we aim at designing filters, digital filters. And for all multirate structures like the one shown in this particular diagram, we always aim at designing filter banks. So, we have low pass filters and high pass filters, they are duals of each other, and they together formulate filter bank. And that is the whole point, when it comes to design of a multirate structure.

(Refer Slide Time: 05:20)



We also typically saw, that if we have a given input signal p of n that belongs to V_1 . Then the virtue of having these disseminators, we end up with projections of the same signal let us say p_0 , that belongs to V_0 , because we are talking about low pass filter, the scaling function, and correspondingly q_0 n W_0 , because we are talking about the high pass filter, which comes out of the wavelet function. And then if we see the ladder structure once again, then we realized that I can draw this ladder structure like this.

(Refer Slide Time: 06:06)



V of minus 3, just contained V of minus 2, and this will continue. And if we are starting in V of 0, not necessarily, we have to start in V of 0, this is just one example. If we start in V of 0, then by virtue of having a two band filter bank structure, we in a way move down the ladder. And typically it is easier to move down the ladder, because if I have the information, and if that information is projected in a subspace V of 0, then what I can do is, I can throw away the information which is W 0, and what I will be left with is V of minus 1. Similarly, from V of minus 1, I can throw away the information, which is W of minus 1, and I can have the information projections in V of minus 2. So, when it comes to moving down the ladder, you have the piece of information and just keep on losing the details, and you get the approximations, in the lower scales.

However, we also realized, that for many of the applications, it is desired it is required to actually move up the ladder. And if I really want to move up the ladder, for example if I have to move from V of 0 to V of 1, then I have to have this information with me. I can add these two pieces and only then I can move to V of 1. From V of 1 if I have to move to V of 2, then I have to have the corresponding projections in W of 1, and these two together can be added, to produce the corresponding projections in V of 2. So, how to go about doing this, and last lecture we in a way put down the framework for doing the same, and what we saw in the last lecture was this.

(Refer Slide Time: 08:42)

$$\begin{aligned} f_j(x) &\in V_j, \text{ Scale} = \frac{1}{2^j} \\ \{ \underline{2^{j/2}} \phi(2^j x - k) \}_k \\ &\Downarrow \\ &\text{Normalization} \\ f_j(x) &= \sum_k \alpha_{j,k} 2^{j/2} \phi(2^j x - k) \\ \alpha_{j,k} &= \int_{-\infty}^{\infty} f_j(x) 2^{j/2} \phi(2^j x - k) \cdot dx \end{aligned}$$

If I have some signal, let us say f_j of x that belongs to V of j . And for the scale of 1 upon 2 to the power j , I can span the whole space, whole linear space using this kind of basis function, where 2 to the power j by 2 is the normalizing factor, and by adding this factor we not only make sure that my basis functions are orthogonal, but I normalize my orthogonal basis function, and such kind of basis functions are indeed orthonormal basis functions. And then ϕ of 2 to the power j x minus k , where k is the translation parameter. So, this gives me normalization. This parameter over here will give me scaling, and k is obviously, meant for translation. And by using this assembly, then I can represent f_j of x like this, where all the $\alpha_{j,k}$ values are the approximated values, which can be calculated using this formula.

(Refer Slide Time: 10:02)

$$\begin{aligned} g_j(x) &\in W_j, \text{Scale} = 1/2^j, \\ &\{ 2^{j/2} \psi(2^j x - k) \}_k \\ g_j(x) &= \sum_k \beta_{j,k} 2^{j/2} \psi(2^j x - k). \\ \beta_{j,k} &= \int_{-\infty}^{\infty} g_j(x) \cdot 2^{j/2} \psi(2^j x - k) \cdot dx \end{aligned}$$

We went ahead and we also looked at the framework, for finding on projections in W_j subspaces. Let us say these are named as functions g_j of x , just for the sake of nomenclature we are talking about the same function, which belongs to a subspace W_j . And with the scale of one upon 2 to the power j again, I can span these linear subspaces using this kind of orthonormal basis function, 2 to the power j by 2 psi, which is my wavelet function, 2 to the power j x minus k , where k is again the translation parameter. And I can write down g_j of x like this, where beta of j k are indeed the details, present in the underlying signal, and beta j of k can be calculated like this.

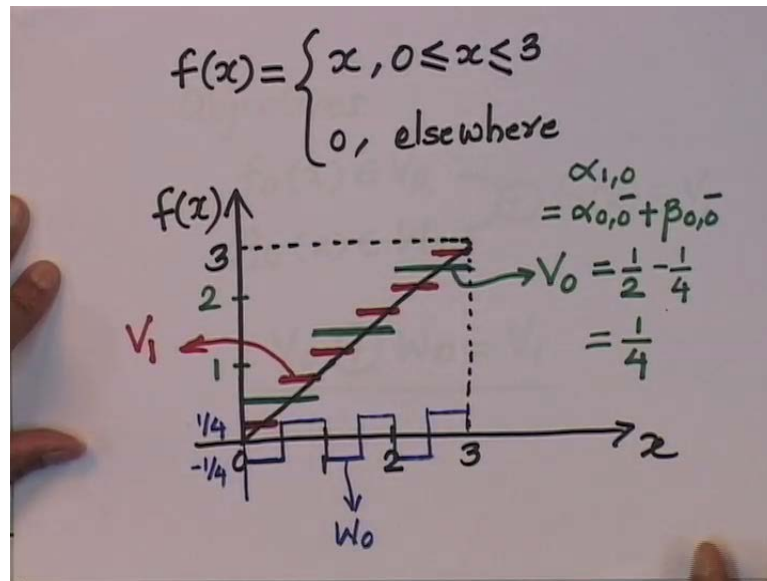
(Refer Slide Time: 11:03)

$$\begin{aligned} \text{Objectives:} \\ f_0(x) &\in V_0 \\ g_0(x) &\in W_0 \end{aligned} \quad \bigoplus \quad \begin{aligned} f_1(x) &\in V_1 \end{aligned}$$

$$V_0 \oplus W_0 = V_1$$

And then we wanted to really ensure that this assembly helps us move up the ladder. If I have projections in V_0 and W_0 , and if I orthogonally add these projections, then I end up with projections in V_1 . So, from V_0 , my journey takes me to V_1 . So, I have start moving up the ladder. And we use this orthogonal summation formula for doing the same.

(Refer Slide Time: 11:36)



If you remember, last time we solved one problem. The problem where we took f of x as the sample function, where x , f of x is equal to x for all the values of x between 0 and 3, and it is a simple straight line. And then we started finding out the projections of this f of x , first of all in V_0 .

(Refer Slide Time: 11:58)

$$\begin{aligned}f_0(x) &= \sum_{k=0}^2 \alpha_{0,k} \phi(x-k) \\ \alpha_{0,0} &= \int_{-\infty}^{\infty} f_0(x) \cdot 2^{j/2} \phi(2^j x - k) \cdot dx \\ &= \int_0^1 x \cdot 1 \cdot dx = \left. \frac{x^2}{2} \right|_0^1 \\ \alpha_{0,0} &= \frac{1}{2}\end{aligned}$$

We know the formula and it all boils down to finding out the alpha values, so alpha 0 0, we plugged in everything, and we realize that alpha 0 0 was half. We went ahead and we also calculated alpha 0 1, and then we realized that alpha 0 1 value came out to be 3 by 2, and correspondingly alpha 0 2 value came out to be 5 by 2. And as the name suggests, these are indeed the approximations, which are shown in green. So, all these green values are the approximated values of the original signal, it is corresponding projections in V_0 using Haar scaling function.

(Refer Slide Time: 12:51)

$$\begin{aligned}\alpha_{0,1} &= \int_1^2 x \cdot \phi(x-1) \cdot dx \\ &= \int_1^2 x \cdot 1 \cdot dx = \left. \frac{x^2}{2} \right|_1^2 \\ \alpha_{0,1} &= \frac{2}{2} - \frac{1}{2} = \frac{3}{2} \\ \alpha_{0,2} &= \int_2^3 x \cdot 1 \cdot dx = \frac{5}{2}\end{aligned}$$

We then went ahead and did the similar exercise, for finding out the projections in W_0 .

(Refer Slide Time: 13:03)

Handwritten mathematical derivation on a whiteboard:

$$g_0(x) \in W_0, \text{ Scale} = 1$$

$$\{ 2^{j/2} \psi(2^j x - k) \}_k$$

$$\{ \psi(x - k) \}_k$$

$$g_0(x) = \sum_{k=0}^2 \beta_{0,k} \psi(x - k)$$

And then we realized that g_0 of x which belongs to W_0 , is indeed summation over the details and the corresponding translated version of the basis function.

(Refer Slide Time: 13:20)

Handwritten mathematical derivation on a whiteboard:

$$\beta_{0,0} = \int_{-\infty}^{\infty} g_0(x) \cdot \psi(x - \overset{\uparrow 0}{k}) \cdot dx$$

$$= \int_{-\infty}^{\infty} x \cdot \psi(x) \cdot dx$$

$$= \int_0^{1/2} x \cdot 1 \cdot dx + \int_{1/2}^1 x \cdot (-1) \cdot dx$$

$$= \left. \frac{x^2}{2} \right|_0^{1/2} - \left. \frac{x^2}{2} \right|_{1/2}^1$$

So, the whole task was, when to find out the beta values. And then we did that, we initially calculated $\beta_{0,0}$, we went ahead and calculated $\beta_{0,1}$, and also $\beta_{0,2}$.

(Refer Slide Time: 13:35)

$$\beta_{0,2} = -\frac{1}{4}$$

$$g_0(x) = \beta_{0,0} \psi(x) + \beta_{0,1} \psi(x-1) + \beta_{0,2} \psi(x-2).$$

And then we realized that all beta values, were indeed minus 1 by 4, and if I have these beta values with me, I can correspondingly find out g_0 of x .

(Refer Slide Time: 13:50)

$$\underline{j=1} \quad f_1(x) \in V_1, \text{ scale} = \frac{1}{2^j} = \frac{1}{2}$$

$$\{ 2^{j/2} \phi(2^j x - k) \}_k$$

$$\{ \sqrt{2} \phi(2x - k) \}_k$$

$$f_1(x) = \sum_{k=0}^5 \alpha_{1,k} \sqrt{2} \phi(2x - k)$$

Then came the important task of moving from V_0 to V_1 . We can go on adding up the alpha 0 and beta 0 k values, to find out alpha 1 k values. We can also calculate alpha 1 k values using the normal formula.

(Refer Slide Time: 14:15)

$$\begin{aligned}
 \alpha_{1,0} &= \int_{-\infty}^{\infty} f_1(x) \cdot 2^{j/2} \phi(2^j x - k) \cdot dx \\
 &= \int_0^{1/2} x \cdot \sqrt{2} \phi(2x) \cdot dx \\
 &= \sqrt{2} \int_0^{1/2} x \cdot 1 \cdot dx \\
 &= \sqrt{2} \cdot \left. \frac{x^2}{2} \right|_0^{1/2} = \sqrt{2} \cdot \frac{1}{8} \\
 \alpha_{1,0} &= \frac{1}{4\sqrt{2}}
 \end{aligned}$$

And in the end, we are able to tally all the different values of alpha 1 k. We solved the problem of alpha 1 0, and we realized that we do not really have to worry about this particular factor of square root of 2, because when it comes to plugging in the value of alpha 1 0 in the formula for f 1 of x, it will be taken care of, and so we should focus only on 1 by four. And then we calculated all the alpha values by adding up alpha 0 k and beta 0 k values. So, we added up all the alpha 0 k and beta 0 k values. For example, in order to calculate alpha 1 0, we added up the first half of alpha 0 0 and the first half of beta 0 0, and that is how we are able to calculate the value of alpha 1 0, and then we continued for the rest of the values, and that is how we are able to generate all these values of alpha 1 k.

We plotted everything and then finally, all the red points, all the red approximated values are indeed alpha 1 k values, and they will help us find out the projections of f of x in V of 1. And we confirmed that indeed we are able to move from V 0 to V of 1, and visually, graphically, we can now verify that V of 1 gives us better approximation of the original signal, when we compare it with the corresponding approximations given by V of 0. So, we are indeed able to add these details in W 0, and then moved on to V of 1, definitely a better approximation of the underlying signal; that is a reason, as to why we revisited this entire example that we solved in the last lecture. What we are going to do in today's class, is in a way based on this particular example that we solved. We are going to pose few interesting questions, based on this particular example.

And then correspondingly try and solve these two questions, and that will formulate the reminder of today's lecture, but before that, let us quickly go through a matlab simulation of the same problem that we just solved. This is a very simple matlab program; `clc` is meant for clearing up the command prompt, `clear all` is going to clear all the variables, and `close all` is going to close everything. Now let us define our x axis, and this could be our time axis for example, and since my signal f of x exists between 0 and 3, x of 0 is 0 and x 1 is 3. The level is selected to be 9 and this level gives us the resolution step, while moving from x 0 to x 1, and that will be clear when we see the next line written over here. So, we have written that x is equal to x of 0. So, it begins with x of 0, and then it increments in a step size, and this step size is $1 \text{ upon } 2 \text{ to the power level}$.

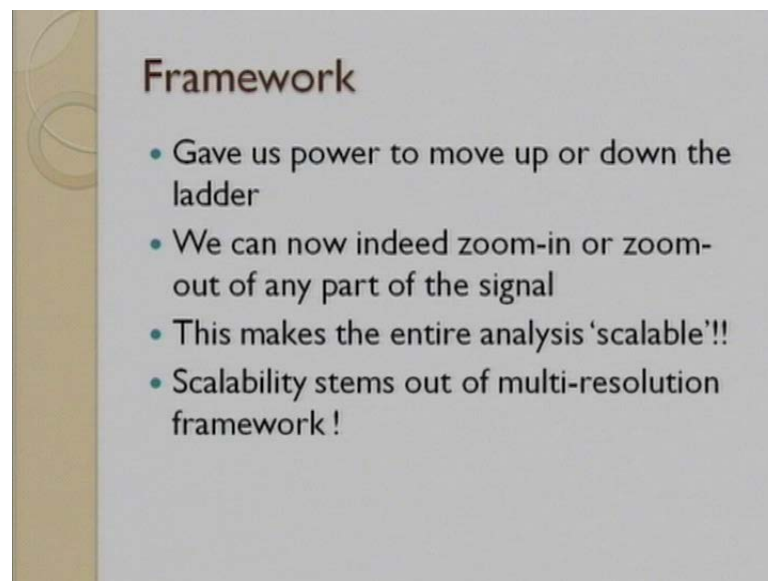
So, the step size that we have selected is dependent on the value of level variable that we select. In this step size, it will start from x 0, and with this step size it will go all the way till x 1, which is 3, minus one of the steps; that is $1 \text{ upon } 2 \text{ to the power level}$. And the underlying function is f of x is equal to x and; that is why we have written f is equal to x , and then we have saved this variable f in a mat file, which is `test underscore f dot mat`. So, let me run this program, and after running this program, we can very clearly see the mat file; that is `test underscore f dot mat`, that has been generated. Now let us invoke the wavelet tool box, and let us look at the one dimensional analysis. Now for this analysis, let us provide the signal that we have generated as the input; that is `test underscore f dot mat`, so this is how our signal actually looks.

And let us tries and analyzes this signal, using a simple Haar wavelet, and let us does the analysis till the third level. So, this is how the analysis actually looks. And now let me be interested in the different mods; that can help us understand what exactly goes into the system. So, these are the separate mods of doing the analysis, and as you can absolutely clearly see, as far as the approximation in a 5 is concerned, it is quite crude. And as we start moving towards a one it gets better and better and better, and we also know the reason, we add a 5 into `d 5` to generate a 4 for example, and this would continue. Let me super impose these modes to really understand it better, and instead of using 5, let me use 10 scales and then try and super impose everything on top of one another.

And now, this particular plot will really help us understand what kind of things we are actually doing, and let us start one step at a time. So, instead of selecting all, let me start with only the most crudest level approximation. And now let us start adding up the

details by virtue of which, we will start moving up the ladder. So, you can clearly see, that approximation is better now, and as you continue doing this, you will get a better and better and better approximation. And then right at the end you go tantalizingly close to your real signal. So, if we just place the last value that we have obtained, then this is indeed very close in terms of representation, to the underlying function, or the underlying signal. Excellent; so, as can be seen from the slides, this particular framework indeed gave us the power to move either down the ladder or up the ladder.

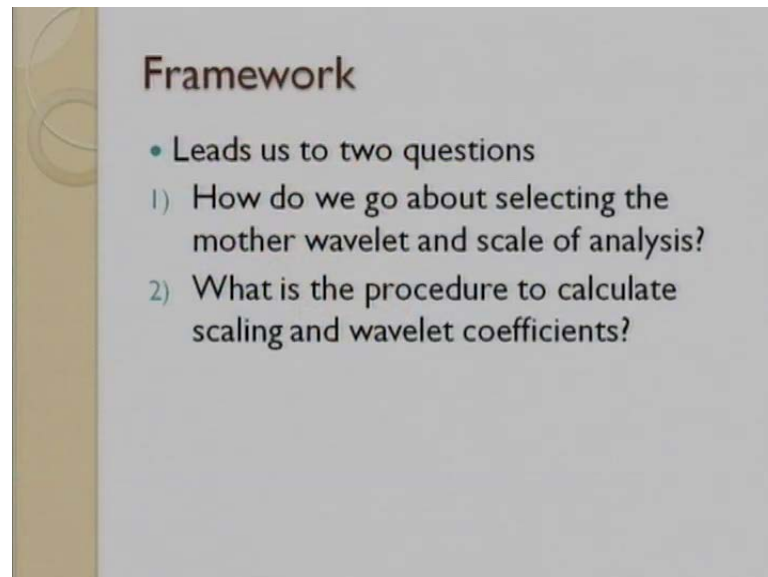
(Refer Slide Time: 22:25)



And we can clearly see; that the zoom in and the zoom out properties of the wavelet transform are indeed met. I can really zoom in and zoom out, and in order to be able to achieve this, I have not only my scaling parameter, but also the translation parameter. The best part is, this makes the entire framework extremely scalable. Scalability stems out of, obviously the whole framework being multiresolution framework. And we make use of scalable systems every now and then, for example all the images which are stored on internet, they are in a way stored in a pyramidal structure, when we Google search for a particular type of image, what gets retrained back, are only the thumbness; that is the upper most slice of the pyramid, and when you select or click on that particular thumbnail, then probably one of the slice, from the underlying framework, of the entire pyramid, will get selected depending upon the resolution of the kind of hardware that you are using. Same holds true even for video analysis, almost all the video codex that we make use of these days, they demand scalability, and that is where wavelets have

started replacing many of the conventional b c structures, for example, in spec discreet cosine transform use to get used. In spec 2000 will all know, we make use of 5 by 3 biorthogonal tab, which is a wavelet basis function, and so scalability is indeed the need of the time.

(Refer Slide Time: 24:46)

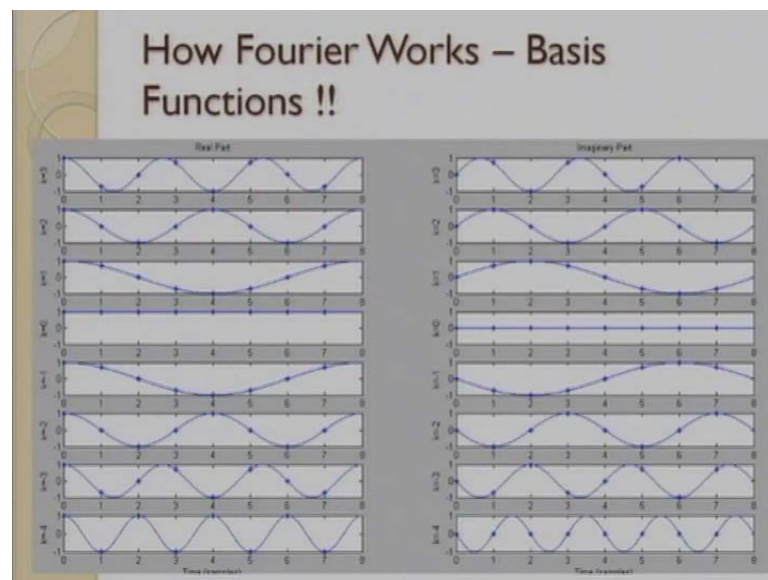


Let us go back to slides, and we have seen the framework, we have studied the framework, and this framework in a way leads us to two very important questions. Question number one, how do we go about selecting the mother wavelet and the scale of analysis. So, both the questions are very important, which are the sub questions of the first question, which mother wavelet, whether I should use Haar mother wavelet or daubechies or shannon or Meyers, there is there is tremendous amount of variety, and what scale of analysis I should select. Should I stop after two scales or three scales or four scales, what is this stopping criteria? These are two important questions. And correspondingly, there comes the second important question as well, and the second question is, what is the procedure to calculate scaling and wavelet coefficients? We have been using the coefficients which are readily available to us, given by doctor Haar, given by doctor Daubechies. Can we think of finding out our own wavelet, can we think of finding out our own coefficients, for doing the analysis. It is a tall order; it is a very difficult preposition all together.

However, at least we can get into the head of the scientists, who in a way discovered these different mother wavelets, and understand the procedure, and who knows that will need us to finding out something of our own. However, we will do that in the next lecture. In the reminder of this particular lecture, we will try and answer the first question, and if we refer to the slides again, the first question is of great significance. The first question in a way asks us, what is the significance of selecting the mother wavelet? If I have an underlying application, and I know what kind of data or information I am going to deal with, then how I can appropriately select the mother wavelet, in order to be able to carry out the analysis effectively and efficiently; that is the question, that we will try and answer in the reminder of this lecture.

And we are going to invoke one important property of wavelets, which is called as vanishing moments. We are slowly and gradually move towards vanishing moments, through correlation. We are very familiar with correlation and convolution. So, through correlation we will slowly and gradually move towards moments; we will define moments, we will understand the concept of moments getting vanished, and we will also try and understand why this is of great importance, great significance.

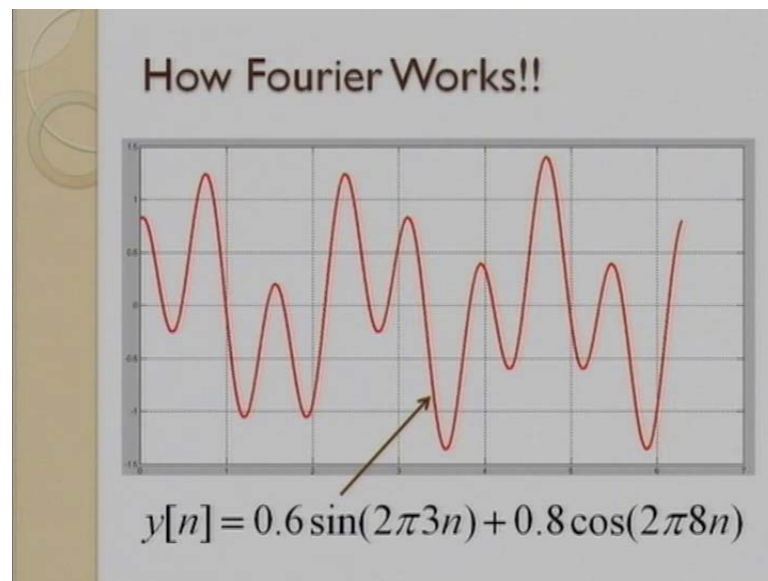
(Refer Slide Time: 28:13)



In order to be able to bring out the correlation parameters in wavelet domain, let us first of all convenience ourselves, that in a way, that is precisely what we were doing as far as the Fourier domain is concerned. In Fourier domain we made use of this correlation to a

large extent, and this particular property of correlation is of great importance, what is shown in this particular slide, these are the basis functions, so typically the real part of the basis, and the imaginary part of the basis. We know the basis function of Fourier transform is $e^{j\omega t}$, and this is the real part, which is the cosine part, and this is the imaginary part, which is the sin part.

(Refer Slide Time: 29:07)



And then, if we pose this particular question, that how indeed Fourier transform works. Then one way of looking at Fourier transform is understanding the correlation between the underlying function and the basis function, and this can be done using simple dot product philosophy. So, what we will do, is we will go back to one sample matlab program. So, in this particular matlab program, as can be seen on the screen, we have generated a function, and we want to analyze this particular function, let us say y . And this function is a good looking stationary function; that means, all the frequencies there present all throughout this function, and as can be seen, in fact there are two frequencies which are present in this particular function; sin of 3 and cos of 8, x is a variable again and you can add a 2π in order to make this normalized. So, this function has 2 frequencies; sin of 3 and cos of 8. And let us find out the dot product.

Well, the concept of dot product is very useful, particularly for engineers while doing the analysis of signals, because in a way dot product means instantaneous product. We can look at all the instances on x axis, which is time axis, and carry out the multiplication as

simple as that. So, keeping that in mind, let us say to analyze this particular underlying signal, I invoke this particular basis in Fourier transform. Now we have understood that Fourier basis is e to the power j , which can be disintegrated into real and imaginary parts, which is cosine and sin part, and let us say I am invoking this underlying signal, with the basis of $\cos$ of $6x$, and if I do so, then let us see what kind of response we get. So, this is my underlying function or a signal, which is stationary signal and it contains two frequencies.

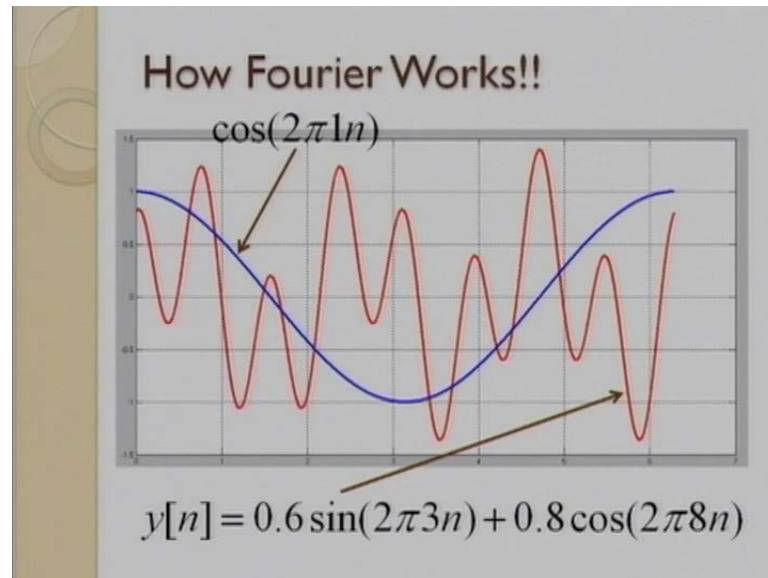
This is my basis function with which I am doing my analysis. So, this is $\cos$ of 6 that is the basis I have used, and you can clearly see, since $\cos$ of 6 is not present in the original signal, there is poor correlation, between the signal to be analyzed which is shown in red and the basis function which is shown in blue. And as a result of that if I do, if I take the dot product between these two then this is what I get. And if I integrate this dot product, which is shown in green, then you can clearly realize that there are as much as positive parts, as much as there are negative parts, and they would cancel out each other, and as a result of that I will not get any peak for a frequency of 6 , and it make sense, because I do not have that frequency in my original signal. So, in my original signal, I have only two frequencies $\sin$ of 3 and $\cos$ of 8 .

So, it make sense, when these parts they would in a way cancel out each other, and that essentially indicates that the underlying signal and basis they are not correlated. Now, instead of using the basis, which is not present in the original signal, using a frequency which is not present in the original signal. Let us invoke the underlying signal with a frequency which is present. For example, $\sin$ of 3 , this frequency is indeed present in the signal, so let us try and invoke with $\sin$ of 3 . And if I run this again, then this is the original signal that we are trying to analyze, and this is now $\sin$ of 3 , and you can clearly observe a good amount of correlation, this is positive cycle and there is lot of activity in positive area, very little activity in the negative area.

This is the negative cycle, and again very little activity in the positive area for the underlying signal. So, the in general correlation is much better, and that is because this frequency is actually present in the original signal. And now if you find out the dot product, then you will realize that there are very less negative parts, and the area under the positive part is much more. So, when I integrate this dot product, I will definitely get some positive value, and that will be the peak at 3 , at a frequency of 3 , when I am trying

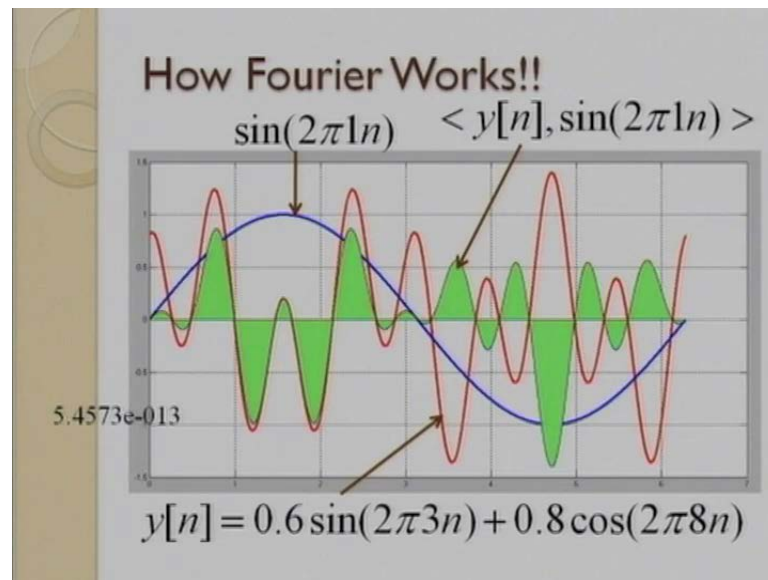
to plot the frequency response or the Fourier response of this underlying signal. So, that is precisely why we end up with a peak, at a frequency of 3. The dot product essentially tells us, that the basis of $\sin 3$ is highly correlated with the underlying signal. This is indeed true for all different varieties of basis. Let us go back to the slides, and let us run through few of the examples quickly.

(Refer Slide Time: 35:43)



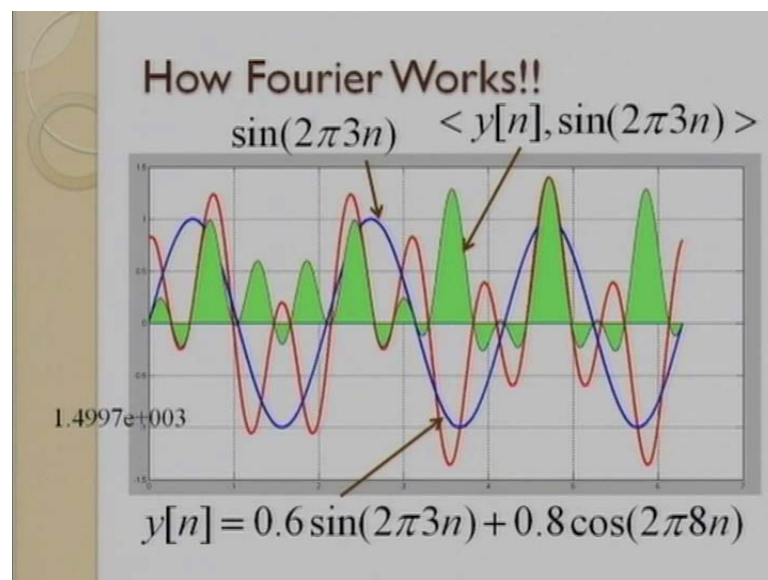
As can be seen in a slide this is the same signal $\sin$ of 3 and $\cos$ of 8, when invoked with $\cos$ of 1, obviously poor correlation, and so the dot product would result into the parts that would in a way cancel out each other. Same signal when invoked with $\sin$ of 1, again this frequency is not present in the original signal.

(Refer Slide Time: 36:07)



So, poor correlation and the dot product parts would cancel out each other. So, no peak will be generated at a frequency of 1.

(Refer Slide Time: 36:19)

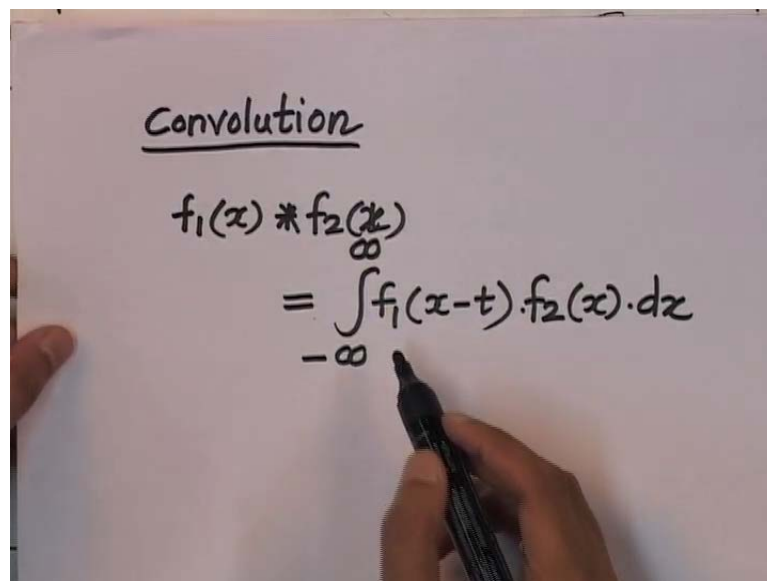


Now if invoke this with cos of 3, you can clearly see that cos of 3 is not present in the original signal. So, once again the dot product will have the parts that would cancel out each other. However, as we have seen in the matlab demonstration, if we invoke this with sin of 3, which is now present in the underlying signal, then definitely dot product

will generate a peak and frequency of 3. The same holds to for cos of 8, because cos of 8 is present in the original signal.

And as a result of that you can clearly sense, that correlation parameter is very high, correlation index is very high, and if you take the dot product, then that is going to give you a peak at a frequency of 8. So, this idea of correlation we have been using with Fourier transform for quite some time. Now this is a good time to invoke the same property of correlation in the context of wavelets. So, let us do that. We know the formula for convolution very well. So, we will start with the convolution formula, move towards correlation formula, and then see what the significance of movements, when it comes to deciding the basis wavelet function, for analyzing a particular type of underlying signal.

(Refer Slide Time: 37:59)



A photograph of a whiteboard with handwritten text. At the top, the word "Convolution" is written and underlined. Below it, the convolution formula is written: $f_1(x) * f_2(x)$ followed by an equals sign and an integral from $-\infty$ to ∞ of $f_1(x-t) \cdot f_2(t) \cdot dt$. A hand holding a black marker is visible at the bottom right, pointing towards the formula.

$$\text{Convolution}$$
$$f_1(x) * f_2(x) = \int_{-\infty}^{\infty} f_1(x-t) \cdot f_2(t) \cdot dt$$

So, let us write down the formula for convolution first, because we are quite familiar with convolution, and if we want to find out the convolution between f_1 of x , convolved with f_2 of x . Then that is going to be equal to integration, from minus infinity to plus infinity, f_1 of x minus t into f_2 of x into dx . Now, if we get rid of this flipping operation, then the convolution will become correlation. So, I can very well write down the correlation formula like this.

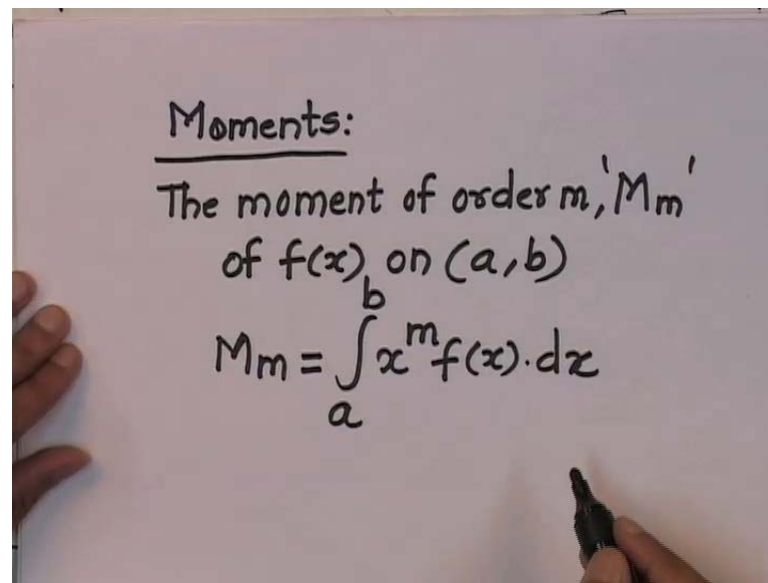
(Refer Slide Time: 39:03)

Correlation:

$$f_1(x) \circledast f_2(x) = \int_{-\infty}^{\infty} f_1(t-x) \cdot f_2(x) \cdot dt$$

I would say correlation between two functions; say f_1 and f_2 , will be f_1 of x . Well this stands for correlation; this is not a very familiar mathematical notation, so please pardon me for abuse of mathematical notation, but just to understand. So, this is correlation between f_1 of x , and f_2 of x and that is equal to integration from minus infinity to plus infinity, I would say f_1 of t minus x now, into f_2 of x , into. We used dx in the last formula. It should be dt , just a correction. So, this is the correlation formula. Now see since we are often interested in detecting sensitive activities in the signal; such as spike or jump discontinuity, and it is higher order derivatives. We would like to know about what kind of wavelets, will be able to see such activities, and indeed it turns out that we can find wavelets, that can detect or see linear or quadratic or higher polynomial structures, and that brings us to the subject on maximum order of vanishing moments. So, let us define, what do we mean by moments first, and then we will understand the significance of moments, when it comes to selecting a particular type of mother wavelet.

(Refer Slide Time: 41:12)



The image shows a whiteboard with handwritten text. At the top, the word 'Moments:' is underlined. Below it, the text reads 'The moment of order m, 'M_m' of f(x) on (a, b)'. The 'b' in '(a, b)' is positioned above the integral's upper limit. The formula $M_m = \int_a^b x^m f(x) \cdot dx$ is written below the text. A hand holding a black marker is visible at the bottom right of the whiteboard.

Moments:
The moment of order m, 'M_m'
of f(x) on (a, b)
$$M_m = \int_a^b x^m f(x) \cdot dx$$

So, moments; the moment of order m, say m of m, of function f of x on interval a comma b, can be given as m to be equal to integration from a to b, x to the power m, f of x into d x. So, this is how I can go about defining moments, and this is of order m. So, I can find out zeroth moment, by plugging in m small m is equal to 0. Correspondingly, I can go about finding out first moment and second moment and third moment and so on. Now, when these moments of higher order they vanish. They in a way tell us something very special about the underlying wavelet. So, what special information these moments convey, we are going to understand in few minutes, but before that to understand whether the moments actually vanish or not. Let us very quickly solve a simple problem with the Haar wavelet once again, to really convince ourselves that the moments indeed vanish out.

(Refer Slide Time: 43:17)

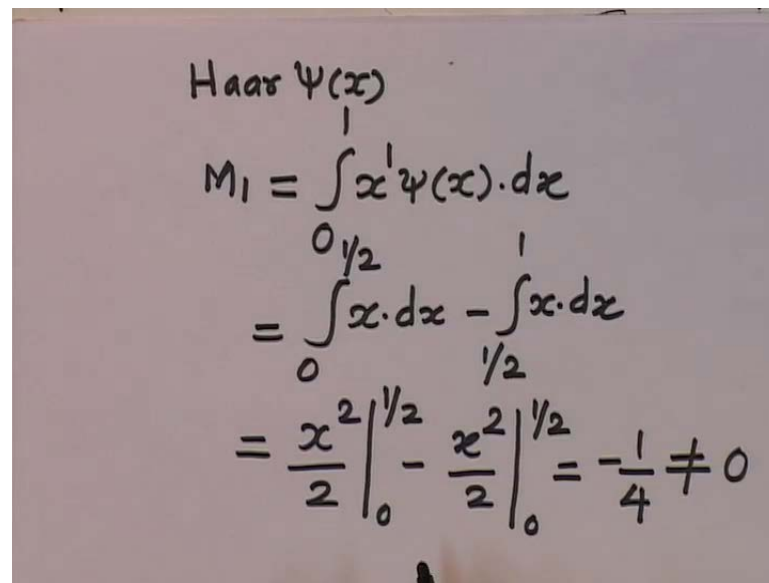
The image shows a handwritten derivation on a piece of paper. At the top, it says 'Haar = Daub1 $\psi(x)$ '. To the right is a small graph of the Haar mother wavelet $\psi(x)$ on a coordinate system. The x-axis is labeled 'x' and has tick marks at 0, 1/2, and 1. The y-axis has tick marks at 1 and -1. The function is 1 for $0 \leq x < 1/2$ and -1 for $1/2 \leq x < 1$. The derivation below the graph is as follows:

$$\begin{aligned} M_0 &= \int_0^1 x^0 \psi(x) \cdot dx \\ &= \int_0^{1/2} 1 \cdot 1 \cdot dx - \int_{1/2}^1 1 \cdot dx \\ &= x \Big|_0^{1/2} - x \Big|_{1/2}^1 = \frac{1}{2} - \frac{1}{2} = 0 \end{aligned}$$

So, let us take a case of Haar, which is also the first member in the daubechies family, so Haar which is equal to daub 1, psi of x. So, let us try and solve for this particular mother wavelet, and let us say I want to find out, what happens to the 0 moment, and we know the formula, and I can plugin the values from 0 to 1; that is because this psi of x exists only between 0 and 1, we typically know how this psi of x looks. So, this is my psi of x, which exists only between 0 and 1, and that is why integration limits are from 0 to 1, and then x to the power 0 psi of x into d x.

And now, we will have to spread this integration into two parts; from 0 to half, we have x to the power 0 is going to go to 1, so 1, into from 0 to half this value, this function takes a value of plus 1. So, I can plug-in 1 of d x minus, because it is going to take a value of minus 1 between half and 1, and then 1 into d x. So, that will result into x 0 to half minus x from half to 1. And so we are talking about a situation, where we have half minus half, and this goes to 0. And so I can very well say that this zeroth moment of Haar mother wavelet vanishes. Whether this holds true for the first moment, as far as Haar mother wavelet is concerned. Let us try and find out.

(Refer Slide Time: 45:32)

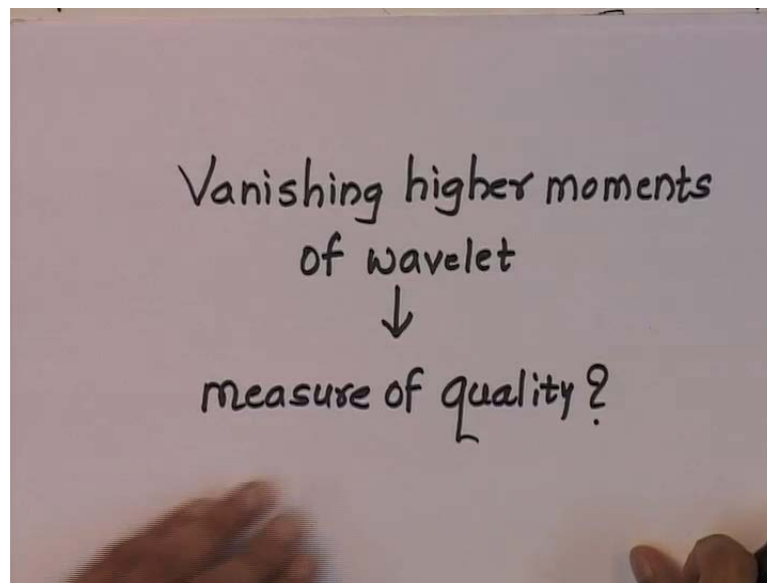


The image shows a handwritten derivation on a chalkboard. At the top, it says 'Haar $\psi(x)$ '. Below that, the first moment M_1 is defined as the integral of $x^1 \psi(x) \cdot dx$ from 0 to 1. This is then split into two integrals: from 0 to 1/2 and from 1/2 to 1. The first integral is $\int_0^{1/2} x \cdot dx$ and the second is $-\int_{1/2}^1 x \cdot dx$. The final result is calculated as $\frac{x^2}{2} \Big|_0^{1/2} - \frac{x^2}{2} \Big|_{1/2}^1 = -\frac{1}{4} \neq 0$.

$$\begin{aligned} \text{Haar } \psi(x) \\ M_1 &= \int_0^1 x^1 \psi(x) \cdot dx \\ &= \int_0^{1/2} x \cdot dx - \int_{1/2}^1 x \cdot dx \\ &= \frac{x^2}{2} \Big|_0^{1/2} - \frac{x^2}{2} \Big|_{1/2}^1 = -\frac{1}{4} \neq 0 \end{aligned}$$

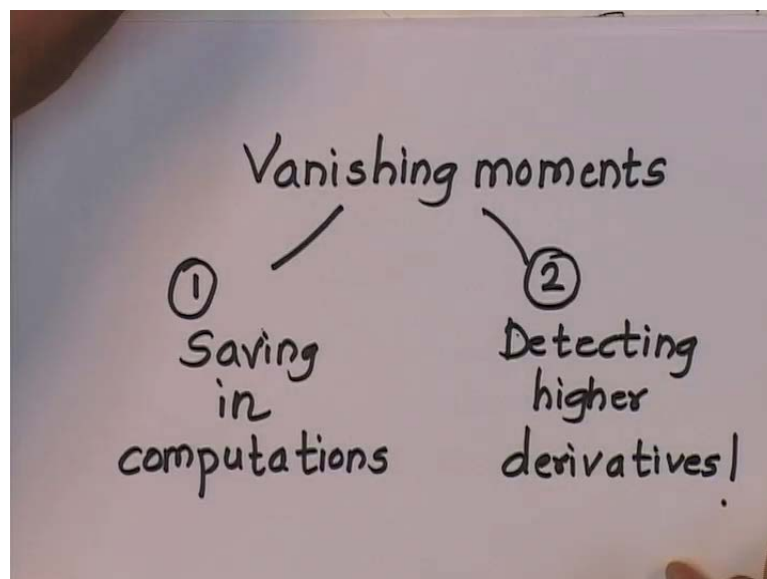
Let us do the same exercise for Haar ψ of x , but let us try and find out the first moment. So, m of 1 is going to be equal to between 0 and 1, and then now I have x to the power 1, ψ of x into dx , and that leads us to from 0 to half, x into dx minus, half to 1, x into dx . and typically we know that this is x square by 2, between 0 and half minus x square by 2 between 0 and half. And if you solve this, this is going to go to minus 1 by 4. It is not equal to 0, and so we can very well conclude, that the first moment of Haar mother wavelet does not vanish, or in other words as far as the Haar mother wavelet is concerned, it has only one moment which vanishes, which is the zeroth moment; interesting.

(Refer Slide Time: 47:05)



Now, that will take us to one very interesting question, and that is vanishing higher moments, and by higher we mean; higher order moments of wavelet, how this is a measure of quality. So, this is one interesting question that we are going to in a way pose. And there are two ways in which we can answer this question, and these two ways are; I can probably say, that as far as my vanishing moments are concerned.

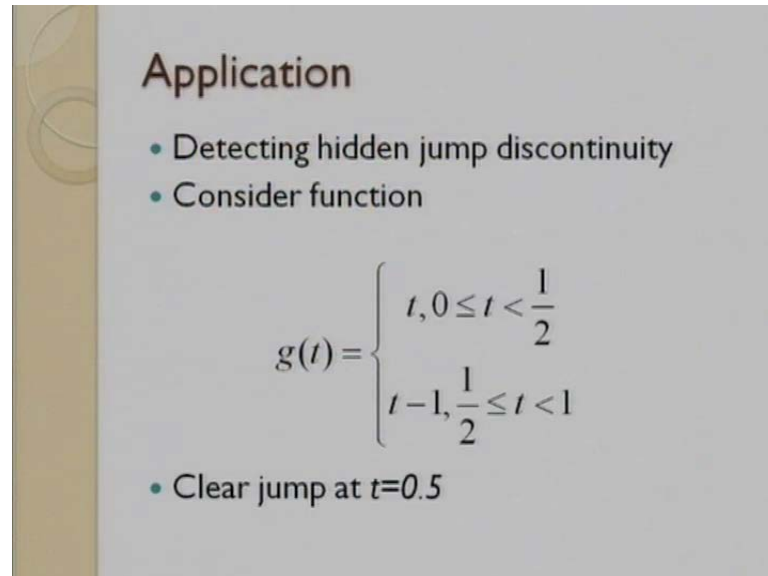
(Refer Slide Time: 48:04)



These vanishing moments, they definitely number one; gives me savings in computations, and number two; they are definitely useful in detecting higher derivatives,

and this second property, is of great importance, it is of great significance. Now, why this is important? Well we are going to revisit the problem that we solved last time.

(Refer Slide Time: 49:10)



Application

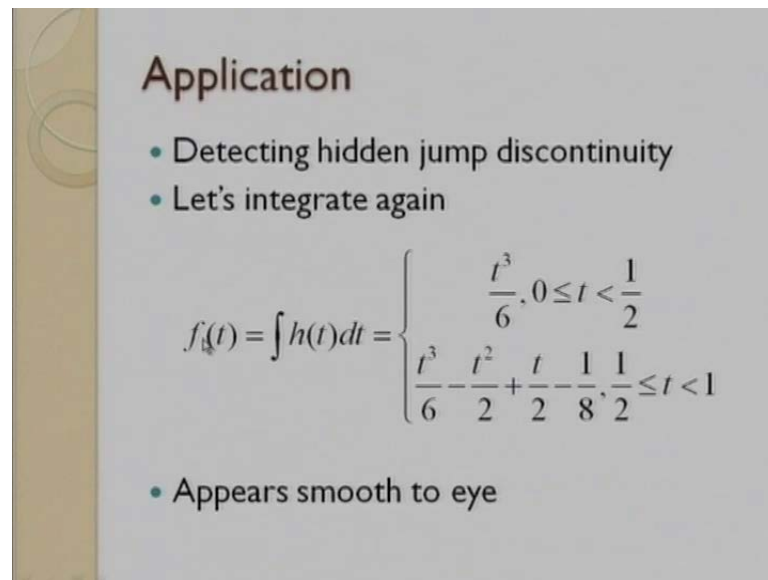
- Detecting hidden jump discontinuity
- Consider function

$$g(t) = \begin{cases} t, & 0 \leq t < \frac{1}{2} \\ t-1, & \frac{1}{2} \leq t < 1 \end{cases}$$

- Clear jump at $t=0.5$

So, we will refer to the slides once again. We solved this particular problem last time if you remember, and this is a very typical application, which is required in many real life situations, where we are trying to in a way find out, the hidden discontinuity, the hidden jump in the underlying signal. And we took this particular function as an example, where g of t is equal to t between 0 and half, and t minus 1 between half and 1, and there was this clear jump at t is equal to 0.5. We did the integration of the original signal g of t , by virtue of doing this integration, we are able to generate signal h of t . And still in h of t , we realize that there is this cusp jump at time t is equal to 0.5.

(Refer Slide Time: 50:15)



Application

- Detecting hidden jump discontinuity
- Let's integrate again

$$f(t) = \int h(t) dt = \begin{cases} \frac{t^3}{6}, & 0 \leq t < \frac{1}{2} \\ \frac{t^3}{6} - \frac{t^2}{2} + \frac{t}{2} - \frac{1}{8}, & \frac{1}{2} \leq t < 1 \end{cases}$$

- Appears smooth to eye

So, we integrated this h of t again in order to finally generate f of t . And now if we plot this f of t , then this function would definitely appear very smooth to our eye. And if you remember, last time we did this small exercise in matlab. So, we are able to generate these different functions, and these different functions they look like this. So, in the original function, the discontinuity is obvious. In the integrated version, there is a cusp discontinuity, and then by virtue of integrating twice we get a very smooth good looking function. And we also save this function, in a test underscore sig underscore f dot mat. And if you remember we did this small exercise last time, that we invoked the windows menu of wavelet toolbox given by math works, and we tried to interpret this particular functions, I am loading this particular signal, that we have generated which is test underscore sig.

And we know that in this particular signal, exactly midway through there is a discontinuity, which is hidden, because we have done double integration of the original signal. So, it is not apparent to the eye, it looks like it is a smooth signal, but we know that there is a hidden discontinuity over here. And now if we try and do the analysis, let us say using Haar, and I am selecting the maximum possible scale. I am going all the way till level number nine, and if I do the analysis and if I see the detail after scale one. Then Haar wavelet analysis misses out this discontinuity. So, I cannot really detect this discontinuity using Haar. By the way Haar is also known as Daubechies 1. So, Haar is

the first member in the Daubechies family. So, Daubechies 1, if I do the analysis then it is the same thing, and we clearly miss out this discontinuity.

As against that if we invoke Daubechies 2, and now if we try and analyze, then you can clearly see that, I can now very clearly find out the discontinuity over here. I can zoom on to this part to really understand it better, and you can clearly see at 256, because we are doing a 2 512 level analysis. So, exactly midway is 256. At 256 this pattern goes abrupt, and it tells us that there is a discontinuity at 256. So, Haar was not able to detect this, but Daubechies 2 was definitely able to detect this, and if you allow me to select Daubechies 3; the discontinuity will be more apparent.

So, it tells us that, there is indeed this discontinuity present at this point in the original signal. Correspondingly, we can go on selecting Daubechies 4 and the discontinuity will become more and more apparent. So, what we are trying to say, is which mother function should be selected, is one very interesting question, and it depends on what kind of data we are trying to analyze. Discontinuity or the hidden jump was detected only using those wavelet functions, which are above Daubechies 2, and there is a strong connection between number of vanishing moments, associated with these different mother wavelets, and their capability in doing the underlying task.

We will stop here as far as this lecture is concerned; and in the next lecture, we will continue happening on the same lines. We will realize that vanishing moments are of great importance, and only certain type and certain kind of wavelet mother wavelets, they are capable of doing certain type of jobs, certain type of tasks. Through the correlation, through the vanishing moments, we will try and make sense of what we covered in this particular lecture that we will conclude in the next lecture. Thank you