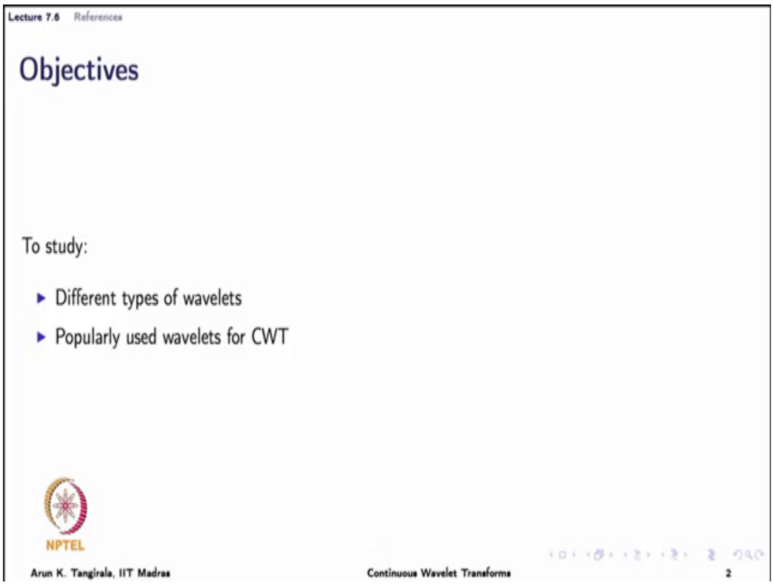
Introduction to Time-Frequency Analysis and Wavelet Transforms

Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

Lecture - 7.6
Wavelets
Part 1/2

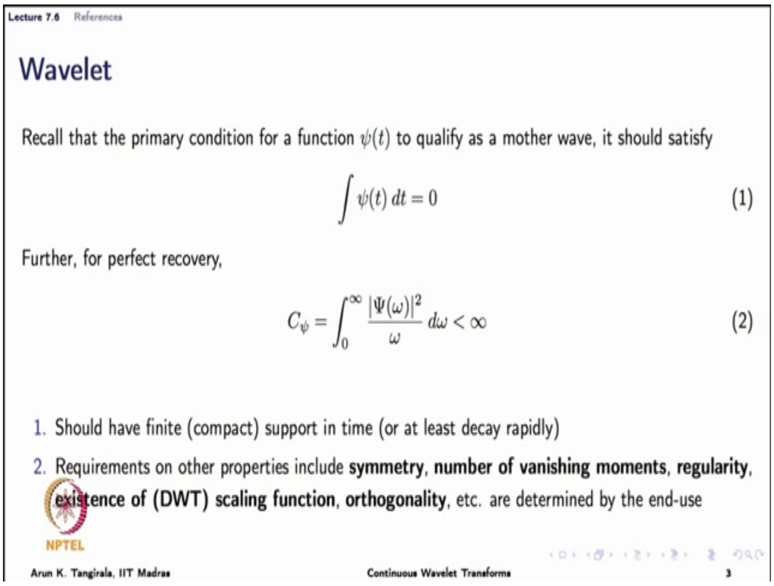
Hello friends, welcome to lecture 7.6 where we shall learn briefly on the different types of wavelets that we use in continuous wavelet transform predominantly. Until now we have learnt the concepts of CWT, the scalogram, and also the concept of scaling function. In this short lecture what we shall do is, we shall look at different types of wavelets.

(Refer Slide Time: 00:49)



Of course, it is not possible to look at all the different types of wavelets that are in the literature but we shall look at a few popularly used wavelets for CWT. And, towards end of the lecture we shall study an important property of a wavelet which is used in selecting the wavelet. This property is known as the vanishing moment's property of the wavelet.

(Refer Slide Time: 01:07)

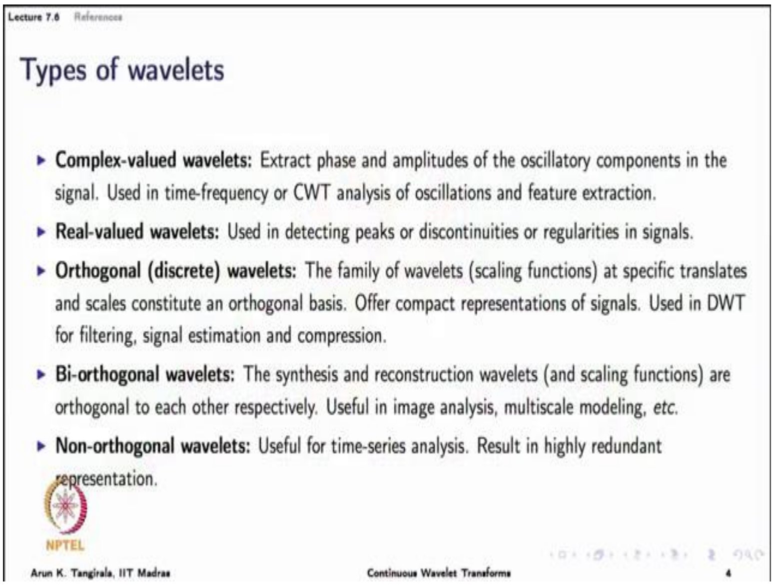


To recap, any function that has a 0 average typically qualifies to be a wavelet. So, you can pick any function that you know which has a 0 average to be used as a wavelet. However, an important requirement is that a wavelet should have compact support; that is it should be localized in time. An alternative way of looking at equation 1 that is a 0 average condition is to say that the wavelet is admissible.

Of course, we have a seen a wavelet before such as a Morlet wavelet which is not really admissible in the sense of 2 that is the, that it does not have a finite admissibility constant. However, we still use that Morlet wave in our wavelet analysis because it decays very fast towards 0 in an effectively finite time. So, in essence, any function that has 0 average from a signal processing view point that is it should act as a band pass filter and one which has a compact support in time qualifies to be a wavelet.

There are other requirements on the wavelet properties such as symmetry, the number of vanishing moments which is something that we will talk about in this lecture; regularity, which is also related to the number of vanishing moments; and, existence of a scaling function in DWT or multi resolution analysis and orthogonality and so on which help us decide or select a wavelet. But, these are not mandatory. These choices are typically driven by the application in hand. Mathematically, it is sufficient to have a function that has finite support in time and that has a 0 average.

(Refer Slide Time: 03:06)



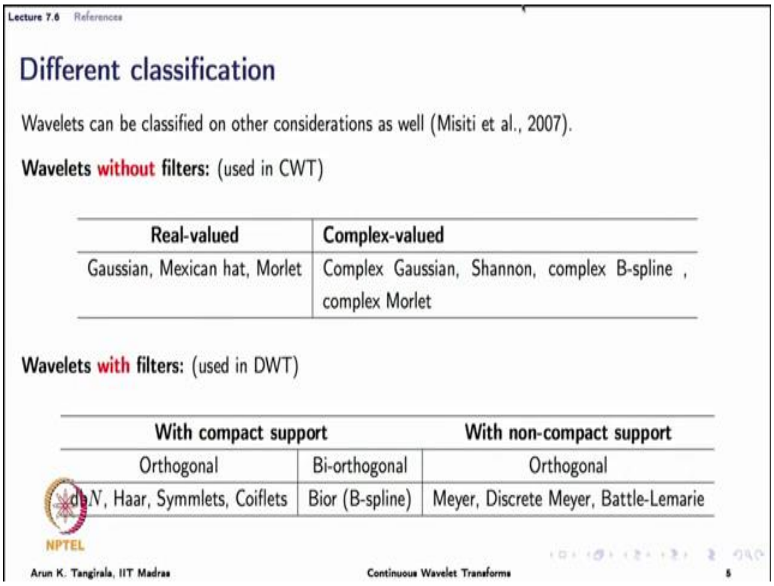
In general, we have different types of wavelets. Obviously, if you look at the conditions that we just stated in saying 1 - 0 average, and 2 - finite support are a very fast decaying time. So, there are obviously a number of functions that would satisfy these conditions. And therefore, you will see number of different wavelets in the literature. And, one can classify this different wavelets based on different criteria. For instance, you could classify them as complex value or real value.

And, this classification is not simply because they are complex valued or real valued, it is also because there is a certain set of applications that each class of this wavelets have. So, for instance, in complex valued, if you take complex valued wavelets they are useful in extracting the phase and amplitude of oscillations. In general, they are used in the time frequency analysis of oscillations or to extract features in the time frequency plane. On the other hand, if you look at real valued wavelets, they are used in detecting peaks or discontinuities; Gaussian derivatives are very useful in this respect.

When you move to discrete wavelets transform you have another set of classification which is based on whether the wavelets are orthogonal to the scaling functions, or we have another classified wavelets as bi-orthogonal wavelets. These terminologies will become clear to us when we learn discrete wavelet transform in detail in the next unit. So, for the moment we will skip the discussion on orthogonal and bi-orthogonal wavelets, but I just given here for completeness sake.

And then, on the same note you have also what are known as non orthogonal wavelets. So, as far as continuous wavelet transform is concerned we are particularly going to look at complex value and real value wavelets. So, those are the 2 classes of wavelets of interest.

(Refer Slide Time: 05:11)



As I said there are different ways of classifying wavelets. Now, here there is a different classification based on whether the wavelets are going to be used in CWT or DWT. Here, I have the, on the top here I have wavelets without filters. Now, this should not convey the message to you that wavelets do not act as filters. What this means is that you cannot really use the standard filtering algorithm that is used in discrete wavelet transform.

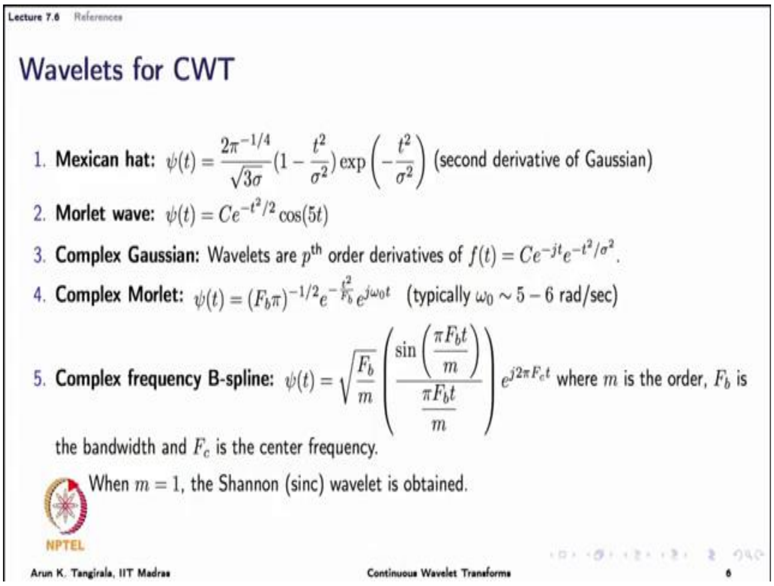
In other words, there are no filters that you can associate directly with wavelets, although these wavelets have filtering characteristics. These wavelets are those that have closed form expressions, analytical expressions, and you can directly use them. Unlike in the case of DWT where typically the wavelets that are used do not have a closed form expression, but a, but the wavelet transforms itself is implemented using the filtering algorithm. So, that is the prime difference.

If you go back to the wavelets that are used in CWT, then again as we have seen earlier we have classification based on whether they are real valued or complex valued. So, I have again some examples for you. In the class of real valued wavelets we would normally encounter what are known as Gaussian wavelets or the derivatives, to be precise the second derivatives of the Gaussian wavelets is the Mexican hat. And, you also have a real valued Morlet wavelet which is something that we have seen in previous lectures, particularly in lecture 7.2 where we looked at scale to frequency.

In the category of complex valued wavelets we have complex Gaussian wavelets, Shannon which is essentially sin c or sinc functions, then you have complex B-spline wavelets which are the generalization of Shannon wavelets, and then you have the familiar complex Morlet wave.

As I pointed out earlier, the real valued wavelets are typically used in singularity deduction which we shall see in the next lecture that is lecture 7.7. And, complex valued wavelets are used in oscillation deduction, in extracting the phase and amplitude of the time varying oscillations which also we shall briefly look at in terms of an application of CWT in the next lecture.

(Refer Slide Time: 07:44)



As I mentioned earlier, we are particularly interested in the wavelets for continuous wavelet transform. And, here I give you the analytical expressions for the wavelets that I have used in continuous wavelet transform. Again, we have the real valued wavelets and the complex versions of the same. For instance, I have the Mexican hat wavelet here; we have looked at this wave.

Although I say wavelet here, these are all mother waves. We have looked at this mother wave when we were discussing the concept of scaling functions. And again, the scaling functions that I have used for CWT are different from those that are used in DWT.

So, coming back to the point here, the Mexican hat wavelets is nothing but the second derivative of Gaussian. In fact, as a simple exercise you should start from a Gaussian function and evaluate the second derivative, and indeed verify that you get this expression. Notice that the constant that appears here, 2 pi raise to minus 1 over 4, over root 3 sigma, is essentially going to ensure that the mother wave has a unit energy. So, the rest of the expression should come out of your second derivative of the Gaussian.

Then you have the Morlet wavelet, real Morlet wave which is essentially the real part of the complex Morlet that you see here. Now, complex Morlet is nothing but an amplitude modulated complex sine wave; as you can see in item number 4, we have a constant multiplied by e to the minus t square by F b times e to the j omega 0 t, omega 0 is a center frequency, F b gives you the bandwidth of the complex Morlet.

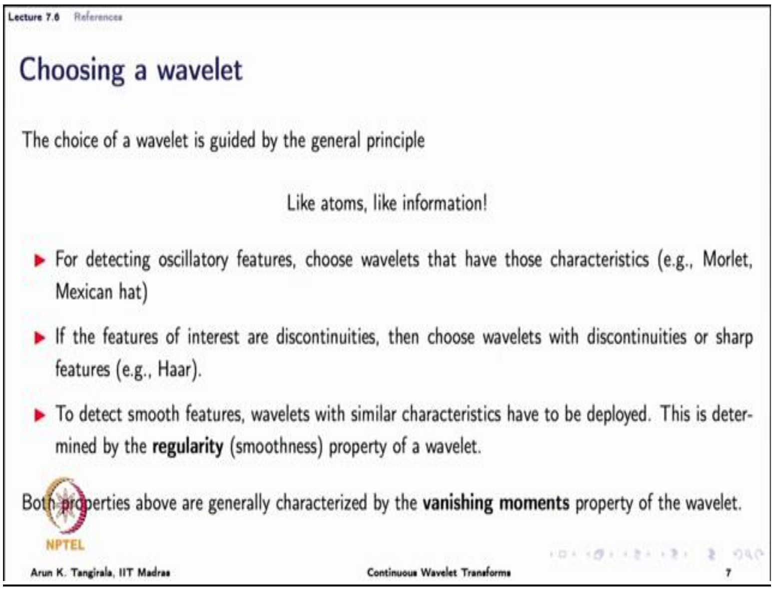
If you look upon this Morlet as a amplitude modulated sine wave then you do understand Morlet, you do understand why this qualifies to be wave, although in a strict mathematical sense it is not a wavelet. The moment I amplitude modulate a sine wave, the existence of sine wave is highly localized in time is restricted effectively to a certain finite time interval. And therefore, atleast in an intuitive sense qualifies to be a wave, but in a strict mathematical sense this is not wave because the admissibility constant for this Morlet wave is as very large that is mathematically infinity.

Now, Morlet used to work in seismology, a lot of seismic data analysis, and therefore, in fact, sorry, not just seismological data analysis but also geo exploration and so on. And therefore, he devised this wave, but later on when the formalizations took place then other waves, wavelets came along which where mathematically qualified to be called as wavelets. So, if you pick the real part of this complex Morlet, I get the Morlet wave here. Again the factor c has to be adjusted such that the real Morlet wave has unit energy.

Then, we have the complex Gaussian which is essentially a complex version of the Mexican hat or the real Gaussian wavelet. Finally, we have the complex frequency B-spline. Now, this complex frequency B-splines are essentially generalizations of the sine c functions, as you can see here. The inner one here, here that you see in the bracket, is a sine c function. And then, of course, you have e to the j 2 pi F c t.

Overall, there are 3 parameters that characterize a complex B-spline. One is the order, and essentially your splines will look like polynomials of a certain order; and, m is the order here; F b is the bandwidth in the, obviously, in the frequency; and F c is a center frequency. When you set n equals 1, you get the sinc wavelet, sine c wavelet. You can obtain plots of all of this with the wavelet tool box and MATLAB, and get a better feel of how they look like. I will show you in this session how to generate different wavelets in MATLAB.

(Refer Slide Time: 12:09)



But before we do that we come to this general question that always one faces in wavelet analysis, how do I choose a wavelet for a particular application? Well, there is no universal answer. It all depends on what kind of information you want to extract, what kind of features are you searching for in signal analysis. And, always remember, like atoms like information.

So, for example, if you want to detect oscillatory features, you should choose wavelets that have those characteristics. For instance, some Morlet wave would have those characteristics because a Morlet wave is an amplitude modulated sine wave, and you have seen many a times how it has an oscillatory nature to it. The Mexican hat wave is also used frequently in oscillation characterization.

On the other hand, a very widely encountered application is a detection of discontinuities or break points you can say or singularities, there are different terms used for the same. Then, you should choose a wavelet that will detect this discontinuity for you which means a wavelet should have discontinuity in it.

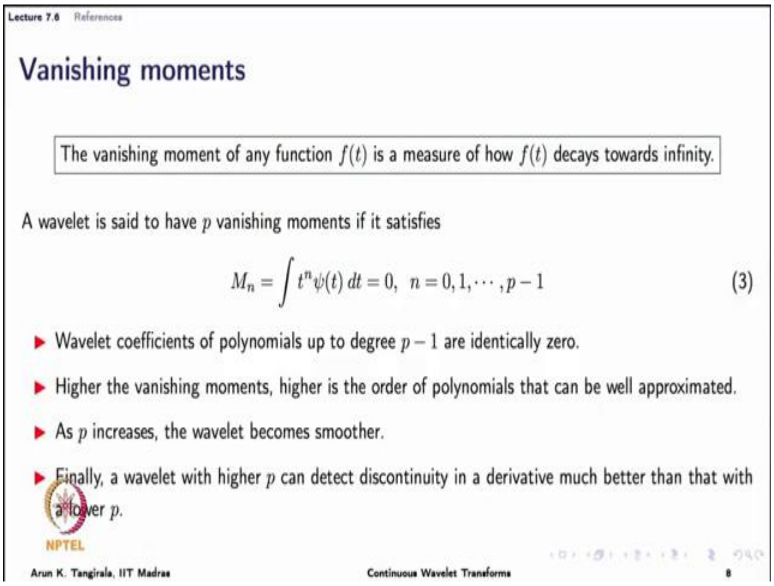
Now, the question is whether you are searching for a singularity in the signal or its derivatives. If you are searching for discontinuity or singularity in the signal itself, then the wavelet itself should have a discontinuity. A very popular wavelet in this category is a Haar wavelet. But, there are many signals in which the discontinuities are not in the signal and they are in the derivatives, first derivative, second derivative, and so on, then you should be choosing a wavelet which also has a similar property.

And, this is also related to detecting smooth features. So, singularities and smooth features complement each other. If I either I want to detect singularities or smooth features of the signal, the technical term for smoothness of any function or signal is regularity. So, if I want to deduct smooth features I should use a wavelet that is so called regular or it has high regularity.

It turns out that both the detection of discontinuities as well as detection of smooth features is characterized by what is known as, that in wavelet analysis, is characterized by what is known as vanishing moments of the wavelet. So, if I know the vanishing moments of the wavelet then I will be able to say whether this wavelet can detect singularities in this signal or its derivatives for instance.

And also, I will be able to say whether a wavelet is suited for deducting smooth features of a certain order because smoothness can be of any order. In fact, there are quantified measured of smoothness such as Folder’s exponent and Richfield’s exponent which tell you how regular the signal is. By knowing the vanishing moments of the wavelets we can relate the vanishing moments to the regularity of the signal and so on.

(Refer Slide Time: 15:18)



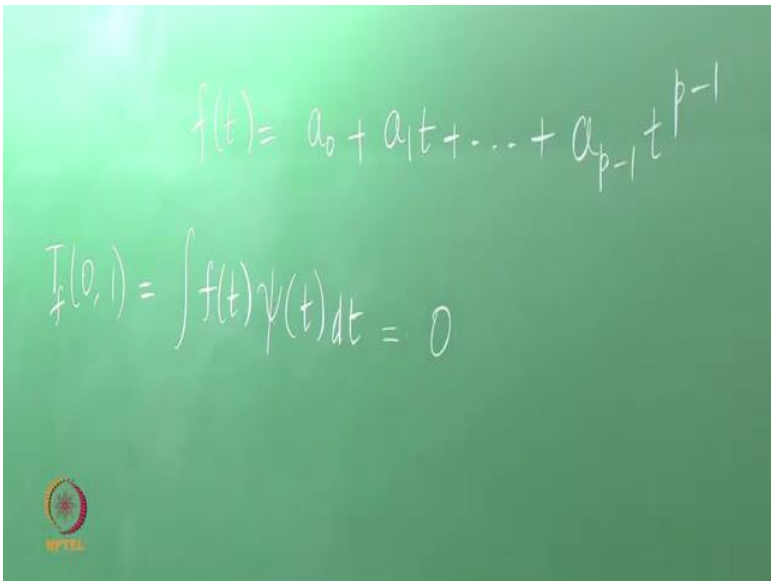
So, let us look at the definition of vanishing moments. First of all, let us remember that the vanishing moments of any function is a measure of how that function decays towards infinity. That is, whether it decays abruptly, whether it decays in an exponential fashion, it decays as a functions of 1 over t, 1 over t square, 1 over t cube, and so on. So, let us look at the mathematical definition of the vanishing moments.

Well, the term itself tells us quite a bit. First we have a moment and then we are talking about vanishing moments. The moment of any function in time or in space and so on, is simply the standard moment definition, integral t power n psi of d t, where psi is the wavelet function that we are looking at. Now, what vanishing moments would mean is that these moments should vanish essentially. So, wavelet is set to have p vanishing moments. If its moments upto p minus 1 are all 0, right, so, therefore, it has p vanishing moments. Notice that the 0th moment is also included in the definition. Therefore, we are looking at p vanishing moments here.

Obviously, any wavelet by definition will satisfy, will have the, will satisfy the equation for n equals 0 because when n equals 0 you are looking at average of the wavelet. Now, the remaining moments whether they vanish or not, depends on the wavelet. We will look at an example shortly, but before we do that let us understand intuitively what this vanishing moments mean for a wavelet.

There are several interpretations that can be given. For example, if a wavelet has p vanishing moments then the wavelet coefficient, that is if I take a wavelet transform and look at the coefficients of polynomials, all polynomials, upto degree p minus 1, they will turn out to be 0. And, that is fairly obvious because if I have a polynomial in time of degree p minus 1 then it would look likes this.

(Refer Slide Time: 17:34)



So, this is a polynomial of degree p minus 1. When I t to wavelet transforms, let us say even with the mother wave, then I have this situation; let us say, that I am looking at a wavelet transform at translation para 0 and scale 1. Then clearly, if the wavelet has p vanishing moments, all of them will turn out to, this wave, the wavelet transform will turn out to be 0 at this scale 1. Because, if just plug in this expression for f (t) here, and then use the property of the wavelet there it has p vanishing moments.

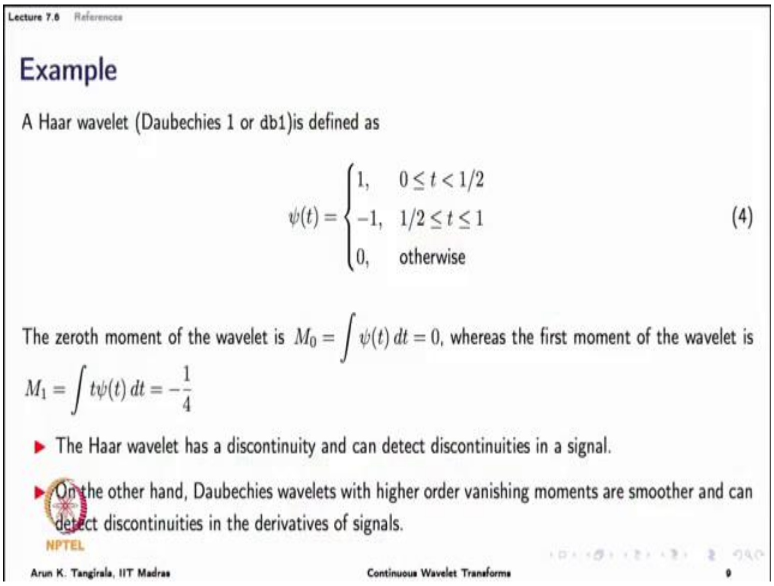
Now, what does that the mean if the wavelet coefficients of polynomials upto degree p minus 1 are identically 0? That means, that higher the vanishing moments, higher the order of polynomials that can be well approximated. Remember, wavelet coefficients are essentially carrying the local details. We are saying all the local details at 0, and the complement of the detail is an approximation, therefore the corresponding scaling function will be able approximate the polynomials of degree upto p minus 1 exactly.

Beyond, if you are looking at a wavelet transform of a polynomial of greater then p minus 1, then the wavelet coefficient would be non 0. And therefore, the approximation of the polynomial with the scaling function is going, not going to be exact. So, that is what it means. And, this has implications in signal compression, as we shall learn in DWT. Because, in signal compression, one of the criteria would be to have as many 0 wavelet coefficients as possible, and as few scaling function coefficients that is as few approximation coefficients as possible.

Therefore, in signal compression we would like to have as many vanishing moments as possible to achieve higher compression factors. But, that will become more clear in DWT. What does it mean for the wavelet itself? So, we have talked about what this vanishing moments means for the wavelet transform. What this means for the wavelet is as p increases that is as the number of vanishing moment of a wavelet increases it becomes smoother and smoother. And, we will look at an example shortly.

Finally, when you are interested in singularity detection, a wavelet with higher number of vanishing moments can detect discontinuities in the higher derivatives. For example, if I have a wavelet with one vanishing moment which is the Haar wavelet, then it can detect discontinuities in the signal itself. But if I look at a Daubechies wavelet, let us say Haar wavelet is also Daubechies wavelet with one vanishing moment, let us say I take a Daubechies wavelet with 4 vanishing moments then it can detect discontinuities in the fourth derivative or the third derivative of a signal, and so on.

(Refer Slide Time: 21:06)



So, let us take an example here of the classical Haar wavelet. We know that the Haar wavelet looks like a box like function which changes, which switches its value after half in the interval 0 to 1. The first moment by definition and even by the definition given in 4 is 0. So, M 0 is 0. We call it as the zeroth moment. Sorry, but this means that we say the wavelet has 1 vanishing moment if wavelet only satisfies M not equals 0.

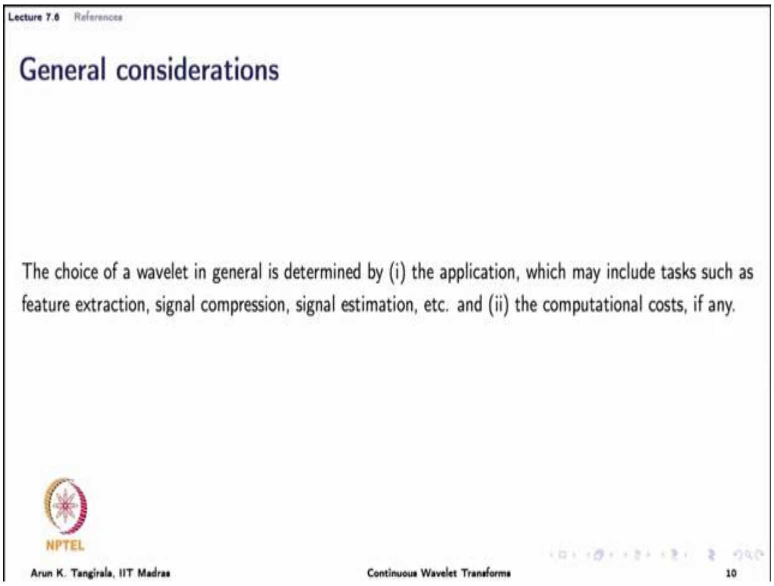
Let us look at the first moment of the wavelet. The first moment of wavelet if you work out the math, so just simply you evaluate this integral, t times psi of t d t with this definition of psi, and simple algebra will tell you, with some integration, will tell you that this value turns out to be minus 1 over 4. You can just take that as an exercise in itself; you can come up with this answer.

Now, what this means is that the first moment of the wavelet does not vanish. In effect, the Haar wavelet has only one vanishing moment; what does it mean for me in signal analysis? It means that it has a discontinuity; it can detect discontinuities in the signals. So, if I want to detect break points in signals, Haar wavelets are good. On the other hand, if you take Daubechies wavelets with higher order vanishing moments, remember that Haar wavelet was shown by Daubechies to be also belonging to the family of Daubechies wavelets with one vanishing moment.

All Daubechies wavelets were designed to have compact support which means they die, they go to 0 exactly in finite time, in Haar wavelet also does. What Daubechies showed is that the existence of, that the higher order vanishing moment of, the vanishing moments property is related to the compact support property, and that we will talk about in DWT.

So, if you move to, if you look at Daubechies wavelets with higher order vanishing moments, they become smoother and smoother and they are more suited for deducting discontinuities in derivatives of signals. So, let us look at that in the MATLAB session, but before we move on to the MATLAB session let us conclude the lecture.

(Refer Slide Time: 23:37)



There are many wavelets available at our disposal. And, we have particularly discussed those that are useful in CWT. We shall look at the other wavelets and their properties for DWT in the next unit where properties such as orthogonality, compact support, bi-orthogonality, and so on, become important. In CWT, primarily we are interested in either detection of oscillations or extractions of features in the time frequency plane, detection of discontinuities or detection of smooth features, these are the typical applications that you would see in CWT.

And, in feature extraction also comes the filtering. Suppose I want to extract a particular feature of the signal then I have to filter, in which case I have to use the inverse CWT. All of these applications that is singularity detection, extracting time features in the time frequency plane, and filtering using inverse CWT, as well as instantaneous frequency detection, we shall learn in the next lecture.

When it comes to choosing a wavelet for CWT, or even in general, the primary consideration is what the application is demanding; whether it is demanding feature extraction, whether it is asking for signal compression or signal estimation and so on, these application signal compression and estimation are typically performed with the help of DWT.

So, the application is the one that is going to really drive the choice of wavelet. And, occasionally the computational cost may matter that is whether the close form expression exists, or how easy it is to compute and so on, whether there exists a fast algorithm. For example, in DWT there exists a fast algorithm. But, you cannot really perform every wavelet analysis with DWT algorithm; you may have to use, you have to come back to CWT for different kinds of analysis. So, we conclude by saying that the choice of a wavelet in general is determined by the application that is induced and the computational costs. And, always remember, like atoms, like information.

(Refer Slide Time: 25:58)

I would strongly recommend that you read these books, by Mallat, of course. But, more importantly the book that I have listed here “Wavelets and their applications” which discuss extensively the properties of each of the wavelets and how they are suited for different applications. Now, let us look at a MATLAB session where I will show you how to generate wavelets and also how to get information on each of these wavelets.