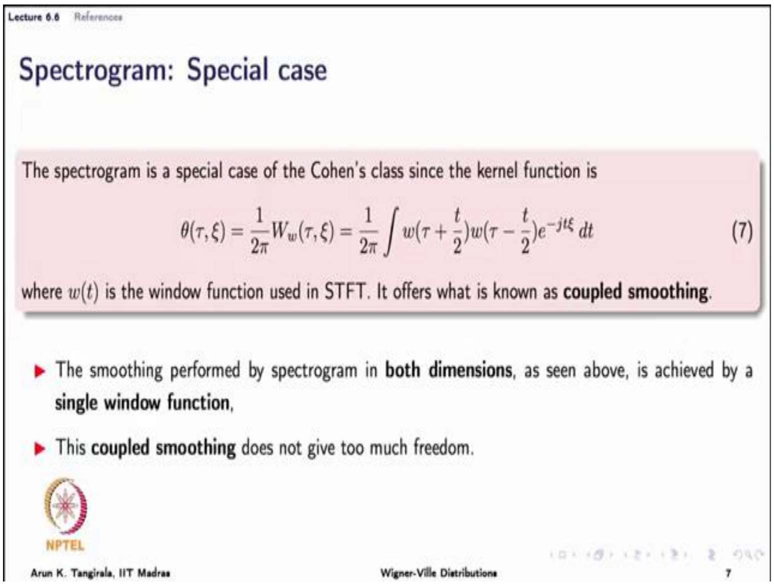
Introduction to Time-Frequency Analysis and Wavelet Transforms
Prof. Arun K. Tangirala
Department of Chemical Engineering
Indian Institute of Technology, Madras

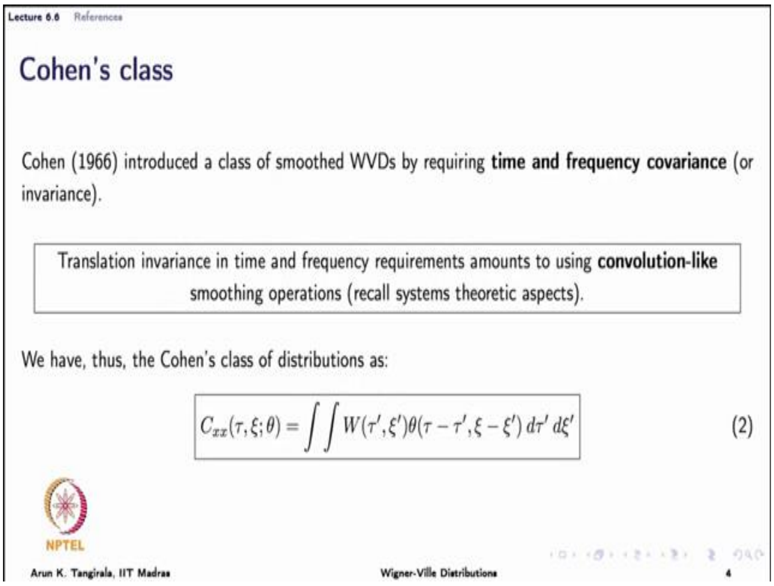
Lecture - 32
Cohen's Class and Smoothed WVD

(Refer Slide Time: 00:15)



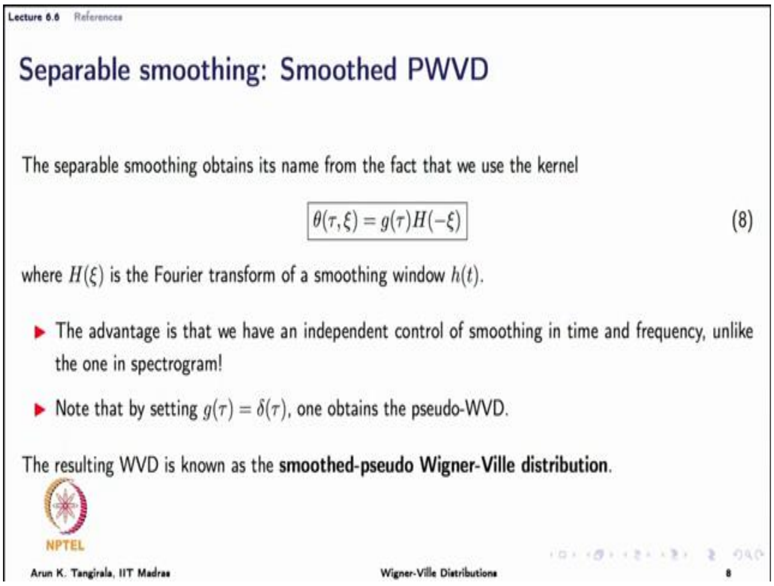
So, now let us discuss why spectrogram is a special case. Well, we know already spectrogram performs convolution operation on the Wigner-Ville distribution. We have seen this equivalence in the previous lecture when we talked about positivity. We showed clearly that the kernel that is used in spectrogram with respect to equation 1; that is with respect to this definition of Cohen's class, we know that the theta there in the case of spectrogram is the Winger-Ville of the window itself that I used in spectrogram.

(Refer Slide Time: 00:37)



So, that is exactly what we have here, but let us reinterpret spectrogram in the context of Cohen's class particularly. What is happening here is the kernel that I have for spectrogram is the Wigner-Ville of the window, and when I look at this expression it is clear that this kernel is surely characterized by single window. What is the role of this kernel? It is supposed to smooth the Wigner-Ville in time and frequency, but I have a single window function which is going to influence how I smooth in time and frequency. That does not give us too much freedom. Ideally when I want to smooth in two dimensions, I would like to have two degrees of freedom; one smoothing operation in time, and another is smoothing operation frequency. They need not be coupled with each other. What spectrogram is doing is, it is actually coupling this smoothing operation in time and frequency domain, and this coupling is facilitated by this window function. They try together in some sense. So, this coupling smoothing may not give us too much freedom. From that view point, we can talk of a separable smoothing in Cohen's class.

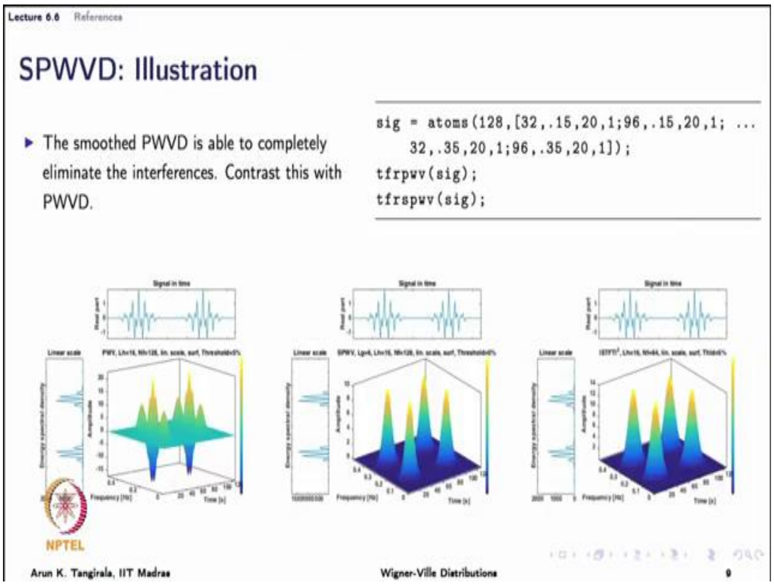
(Refer Slide Time: 02:01)



So, now what we are discussing is slowly the different possibilities for these smoothing kernels in Cohen's class. If I think of a separable smoothing, what you mean by separable smoothing is I would like to have two degrees of freedom when it comes to smoothing. I will perform some kind of averaging in the time along the time access, and another kind of averaging along the frequency access. When I want to do that, I need to introduce a kernel which is separable, because I want that property and a general form for such separable smoothing kernel is given here. It is a product of g of tau times H of minus g. Again here tau and g are dummy variables for time and frequency. So, this g will help me smooth in time and this big H is the Fourier transform of a smoothing window that I will use in time, but which is responsible for frequency. Because I am smoothing in frequency, I look at the Fourier transform of another window function.

So, I am using two windows here. Each window performs smoothing in the respective dimensions. Obviously, the advantages that I have an independent control of smoothing in time and frequency unlike in spectrogram. You should recognize that the movement you set g of tau as delta tau, then you obtain the pseudo Wigner-Ville distribution. I leave that as a simple exercise for u. Because of this that is when I said g of tau as a Dirac that I obtain the pseudo Winger-Ville, these separable smoothing kernel, that is smoothing using the separable kernel is known as a smooth pseudo Wigner-Ville distribution, which means now I have another degree of smoothing in which is happening in the time access. We called the pseudo winger distribution from the previous lecture, we told that it is equivalent to smoothing only along the frequency access. Now, with the help of g of tau, I am introducing smoothing in time as well, and therefore what I have is a smooth pseudo Wigner-Ville distribution. And the smoothing in time is being enabled by g of tau. The smoothing in frequency is enabled by this h, big H or small h which ever they look at.

(Refer Slide Time: 04:21)



So, let us understand with the help of an example the difference between pseudo Wigner-Ville and smooth pseudo Wigner-Ville. So, I have taken the signal that you have seen before in some of the previous lectures. Again I have here two pairs of frequencies, 0.15 and 0.3 at each point. So, I have 0.15, 0.3 here and 0.15 and 0.3 here, but they are located at different time centers, these respective pairs. What you see here on the left most plot is the Wigner-Ville distribution itself. We have the usual spectral plot to the left of this figure. So, the join density here is given by the Wigner-Ville. I am showing you now the 3d version. Until now we are used to looking at the two-dimensional plots where we are looked at contour plots and the color represented the intensity.

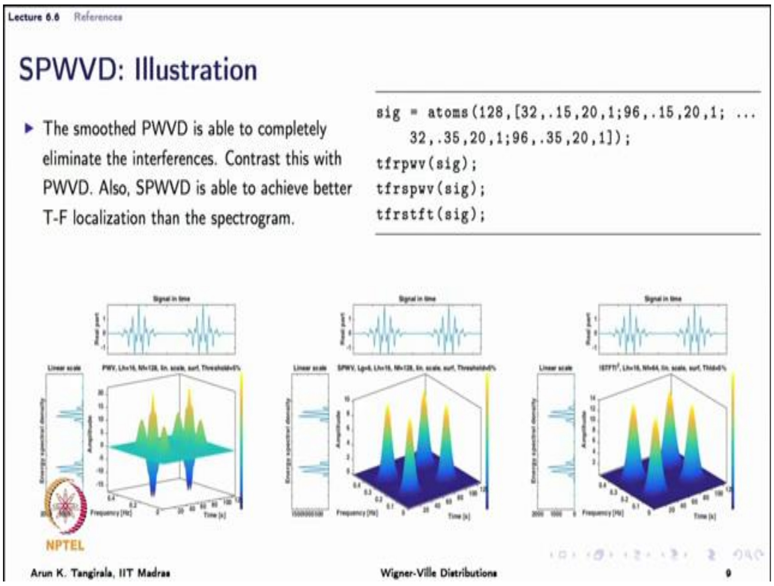
Here the color simply is just done based on the scaling, that is on the values of the amplitude. So, same story here, but I have a three-dimensional plot which gives me (( )) display and information about the distribution itself. The first thing that I note is this Wigner-Ville can take on negative values which I have been talking about time and again. Now, when I perform a smooth Wigner-Ville here, then it is able to completely eliminate the interferences that, sorry I said earlier that this slide allows us to compare smooth pseudo and pseudo Wigner-Ville. It is actually helping us compare smooth Wigner-Ville and smooth pseudo Wigner-Ville and spectrogram.

So, what I have in this center is, I have this smooth Wigner-Ville distribution. It eliminates the interferences completely and on the right most I have this spectrogram. So, what did we say earlier? We said spectrogram is performing some kind of couples smoothing whereas, the smooth pseudo Wigner-Ville gives me independent smoothing in time and frequency. Therefore, I have more degrees of freedom. Now, although it is not so obvious from the 3d plots, you can see that the smooth, the spectrogram as a larger smearing of the energy compared to the smooth pseudo Wigner-Ville. Therefore, this smooth pseudo Wigner-Ville is better here because it achieves positivity; it removes interferences and achieves a better time frequency localization of the energy than the spectrogram.

This again is an illustration of the fact that we have been maintaining right from the beginning is that spectrogram and scalogram are not the only solutions to time frequency analysis. There exists delta quadratic time frequency distributions that could do better than the spectrogram and the scalogram, but one point to keep in mind is that smooth pseudo Wigner-Ville does not necessarily guarantee positivity all the time. It depends on the choice of the window functions and particularly the width of g and h that we use in the time and frequency as a smoothing. If you choose too narrow g and too narrow h, then there is every chance that you will get negative value distribution. So, in order to guarantee positivity, you will have to use wide enough g and wide enough h, but this example tells me they do not have to be as wide as they are in the spectrogram, right. So, you can try some try and error determine the optimum window widths for g and h. There are some default values set in this algorithm t f or on the routine t f of SPWVD.

I think we better take from this slide. The code is also wrong. Yeah I know. I said completely wrong because what I am showing here is pseudo Wigner-Ville. I said we are showing Wigner-Ville [FL]. So, up to this slide it is [FL].

(Refer Slide Time: 09:25)

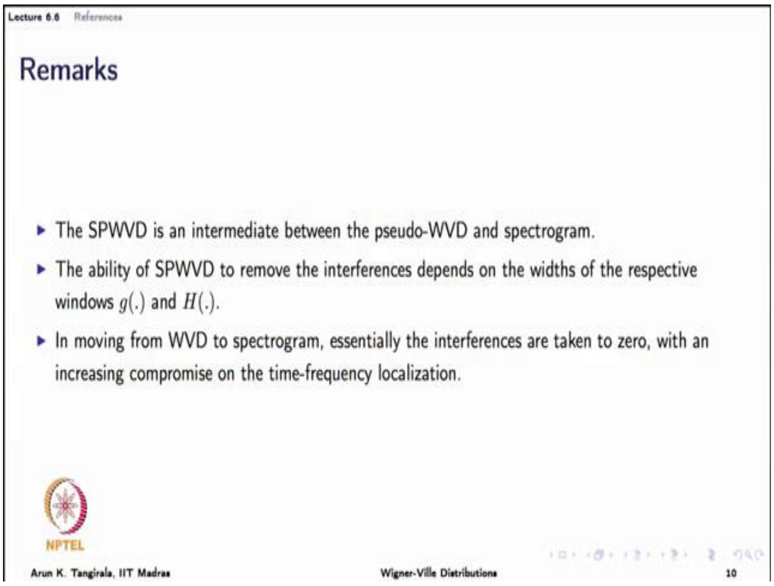


So, let us look at an example to compare the smooth pseudo Wigner-Ville with the pseudo Wigner-Ville and the spectrogram itself. On the left most here, we have the Wigner-Ville of this signal that we have seen earlier, we have a pair of frequencies located at two different time centers. The pseudo Wigner-Ville is an improvement of the Wigner-Ville, but it is not guaranteed to remove the interferences, and therefore you will also have negative valued energy densities in the time frequency, and that is exactly the case here I see in this three-dimensional plot. Earlier we have seen two-dimensional plots, but now for the first time we are looking at three-dimensional plots. It gives much better display, but it is heavy on the graphics sometimes.

So, the pseudo Wigner-Ville gives me negative valued distributions whereas, a smooth pseudo Wigner-Ville is able to get rid of the interferences and by Wigner's theorem we know that positivity is guaranteed and the spectrogram is designed to give me positivity. That is the difference between smooth pseudo Wigner-Ville and the spectrogram. In smooth pseudo Wigner-Ville, there is no explicit requirement that interferences should vanish or positivity should result whereas, in spectrogram the kernel itself is chosen in such a way, and I will give you what the value, what the kernel is in a few minutes. It is guaranteed to generate positive valued distribution. However, notice that the smooth pseudo Wigner-Ville is able to achieve better time frequency localization or smearing than the spectrogram. It may not be so obvious from the three-dimensional plots.

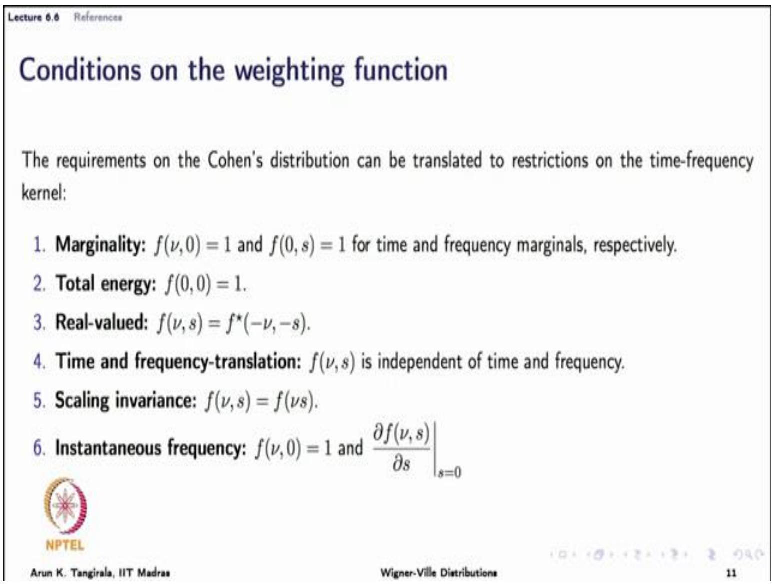
I suggest that you rerun this code and look at the two-dimensional plots. In that way you should remember that the smooth pseudo Wigner-Ville can generate negative value distributions, but if you choose wide enough g and h by some trial and error, you will be able to generate positive valued distribution. Of course, that is taking you closer to spectrogram and there is no doubt about it, but there is an intermediate solution. Spectrogram is only one class of positive valued distributions. Smooth pseudo can also offer another class of positive valued distributions perhaps with a better time frequency localization than what you get in the spectrogram. So, this again is emphasizing a fact that has been maintaining right from the beginning that spectrogram and scalogram are not the only solutions to time frequency analysis. There exists a wealth of other quadratic time frequency distributions that can do a better job than either of them.

(Refer Slide Time: 12:18)



So, these are some of the remarks that we have just talked about. The main message is that the smooth pseudo Wigner-Ville is an intermediate between the pseudo Wigner-Ville and the spectrogram, and this is intermediate that will give you access to repertoire of distributions, right. Now, also the ability of smooth pseudo Wigner-Ville to remove the interferences depends on the widths of the respective windows and in moving from Wigner-Ville to spectrogram, we have taken out the interferences, but we guarantee possibility and that is required by virtual localization.

(Refer Slide Time: 12:50)



So, we shall close this lecture with the quick discussion on the conditions on the weighting function, and also look at just tabulate the kernels which will give rise to well-known distributions. Now, I said earlier on that you can translate the requirements on the distribution to the kernel, and this is exactly in-line with that. If I want marginality, the derivation is… What is important to (( )) that the kernel will determine whether the particular distribution satisfies or possess a certain property.

We shall close this lecture with the discussion of two points. One on the fact that the property of the desire of the distribution that are a desired property of the smooth Wigner-Ville is essentially equivalent to place in certain requirements are achieved by placing certain requirements on the kernel function f itself, and two, we will look at a table of well known distributions set achieved by specific choices of the kernel itself. What you could do is look at the table, look at its requirements and see for yourself what properties each of those distributions satisfy.

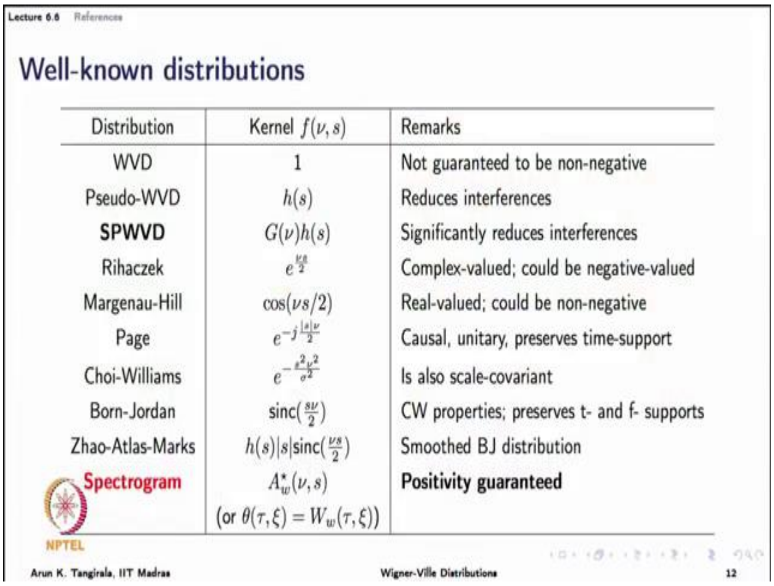
So, let us quickly go through this discussion on the equivalence between the properties of this smooth Wigner-Ville and the restrictions on the kernel. The first property of interest is always marginality, but we know from Wigner's theorems, the moment marginality is satisfied, then there positivity is not guaranteed. So, to ensure that the smooth Wigner-Ville satisfies the marginality requirement, I need to have the kernels satisfy these two requirements that is the value of the function on the respective access should be unity which means that f of nu, 0 equals 1 and f of 0, s should be 1 for time and frequency marginal respectively.

I am not proving this for you. The proof is fairly straight forward. All you have to do is start deriving the expression for the marginal densities from Cohen's class and then ask when the marginality is satisfied and you will come to this requirement. For a detailed derivation, it is not too much to see because it is just about 4 or 5 lines. You can refer to Cohen's book where the restrictions on this kernel for each of these properties are derived in detail.

Now, if I want the total energy to be satisfied that is I integrate I take the area under the smooth Wigner-Ville distribution in the time frequency plane and I want that to be the energy of the signal, then the kernel should necessarily satisfy this property that its value at the origin in the doppler and delay or doppler and shift plane should be 1. If I want real value distributions which Wigner-Ville thus give me, then this kernel should be conjugate symmetric. Likewise if I want time and frequency translations in variance, then the kernel should be independent of time and frequency, that is it should be independent of tau and g and scaling invariance requires that this kernel now should be a function of the product of mu and s. That is wherever mu and s appear, they should always appear as a product in the kernel and we will see Choi Williams distributions is one of those which is scaling invariant.

Recall what is scaling invariance? If I dilate or compress a signal, the same should be reflected in the energy density as well. If I want the instantaneous frequency is to come out right like the Wigner-Ville does, then again the kernel should satisfy these requirements. In fact, you should readily check that the Wigner-Ville satisfies, remember for the satisfaction of all of this and the kernel for the Wigner-Ville is that f is 1 at all points in mu and s. So, it satisfies the total energy, it satisfies the marginality and it satisfies the real valued and so on and the Wigner-Ville has all of these properties, right.

(Refer Slide Time: 17:37)



Distribution	Kernel $f(\nu, s)$	Remarks
WVD	1	Not guaranteed to be non-negative
Pseudo-WVD	$h(s)$	Reduces interferences
SPWVD	$G(\nu)h(s)$	Significantly reduces interferences
Rihaczek	$e^{\frac{\nu s}{2}}$	Complex-valued; could be negative-valued
Margenau-Hill	$\cos(\nu s/2)$	Real-valued; could be non-negative
Page	$e^{-j\frac{\|s\|\nu}{2}}$	Causal, unitary, preserves time-support
Choi-Williams	$e^{-\frac{s^2\nu^2}{\sigma^2}}$	Is also scale-covariant
Born-Jordan	$\text{sinc}(\frac{s\nu}{2})$	CW properties; preserves t- and f- supports
Zhao-Atlas-Marks	$h(s)\|s\|\text{sinc}(\frac{\nu s}{2})$	Smoothed BJ distribution
Spectrogram	$A_w^*(\nu, s)$ (or $\theta(\tau, \xi) = W_w(\tau, \xi)$)	**Positivity guaranteed**

Let us look at this table of some well known distributions which is not exhaustive, but this is some of the most familiar ones. It is natural to start with Wigner-Ville, where this kernel function takes on value of 1 in the entire mu s plane on the doppler delay plane and the problem is it is not guaranteed to be non-negative. Why it is obvious because it satisfies the marginality requirement. It is not guaranteed to be positive by virtue of Wigner's theorems. The pseudo Wigner-Ville uses this kernel h of s. Of course, we have used h of tau earlier and h of t and so on, but remember all of these are dummy variables. Essentially what does it means is pseudo Wigner-Ville performs only smoothing in one dimensions and the smooth pseudo Wigner-Ville is a separable kernel. It actually significantly reduces the interferences.

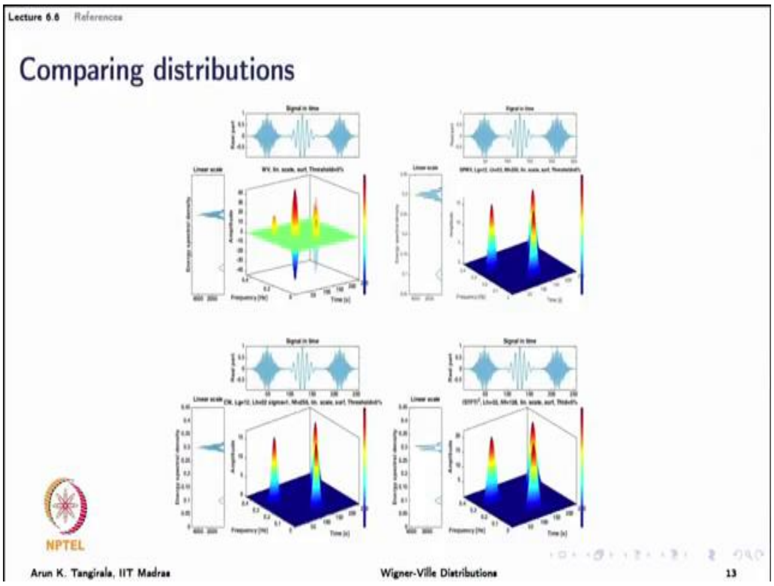
Now, recall that when we talked about smooth pseudo Wigner-Ville, you should not get confused here. The expressions are given for f whereas, if you look at the expression that we gave earlier for smooth pseudo Wigner-Ville, we gave in terms of theta which is a two-dimensional Fourier transform of f. There it was a product of time window and a frequency window because f is the dual of theta, the kernel. Now, the kernel for f is a product of Fourier g of mu and then a product and the time window. So, it is just a dual. That is why you should not get confused between two expressions. They are absolutely equivalent and then you have the familiar Rihaczek distribution which was one of the early distributions to appear in contemporaneously with Wigner-Ville. This is the simple kernel.

The problem with Rihaczek distribution is that it could be complex value. It is also called the negative value, that it has certain simplicity to it. So, it is more of an academic interest to look at this distribution. So, there are other distributions. Let me specifically talk about Choi Williams which I said earlier is scale invariant. Look at the expression for the kernel. We said if I want scale invariance, I want the mu and s to appear in the product in the kernel all the time, and it does appear although this is only single expression. It is a Gaussian like function, Gaussian like kernel in the two-dimensional plane and sigma is a user defined parameter depending on way you choose your sigma. You will the able to get less smearing and more smearing and so on.

So, the advantage of the property of this Choi-Williams distribution also is scaling variant which is not the case with many of the other distributions, of course apart with Wigner-Ville being an exception. Then, there is this Born-Jordan distribution which improves upon the Choi-Williams, where it preserves the time and frequency supports. So, there are certain requirements and many text books list tabulate these requirements on the kernel function, very exhaustive set of properties and you can refer to those books for a complete list of the restrictions on the kernel for properties, and you can see that the Born-Jordan distribution satisfies the requirements on the kernel which will produce time and frequency supports. Then, you have other distribution which is called Zhao Atlas Marks which is a smooth Born-Jordan distribution. So, it performs an additional smoothing.

Of course, the one that is of interest is a spectrogram which once again tells me the spectrogram is a class of smooth Wigner-Ville. Here the kernel is nothing but the ambiguity function of the window. Well, it is a complex conjugate of the ambiguity function of the window. We have not yet studied what ambiguity functions is... In fact, now this profile looks at what are known as ambiguity function. You can also specify remember the kernel in terms of theta which is a dual Fourier, two-dimensional Fourier dual of f. In this case, the theta is the Wigner-Ville of the window itself. This we have seen earlier in the previous lecture and even in this lecture we have shown that spectrogram is smooth Wigner-Ville, where the smoothing kernel or the convolving kernel is the Wigner-Ville of the window itself, and from this you should understand even before we have defined ambiguity function that the ambiguity function and the Wigner-Ville are two-dimensional Fourier transform duals of each other because f and theta are. Therefore, ambiguity and Wigner-Ville should be two-dimensional dual. That is exactly the case and we will see that and will prove that when we talk of ambiguity functions in the next lecture.

(Refer Slide Time: 22:27)



To conclude the lecture, let us take a look at some of these distributions on a signal that we have seen earlier, I have three atoms here. Wave of t of three different frequencies are amplitude modulated. The middle one is the center atom is the one that has the lowest frequency. And what I have shown here is the Wigner-Ville distribution, and then I have shown you the Choi-Williams on the bottom, and on the right top I have this smooth pseudo Wigner-Ville for you with certain choice of windows, I have ensured that these choices will generate positive value distributions. Then I have the spectrogram for you at the right bottom. Obviously, Wigner-Ville gives me the best time frequency localization, that is the energy smearing is minimal, but it generates negative valued distribution. On the other hand, I have spectrogram here which is generating positive valued distributions. Of course, the smooth Wigner-Ville also does, smooth pseudo Wigner-Ville also does that for me and the Choi-Williams also does the same job for me, and you can now pick which distribution is good for this particular sigma.

Once again telling you that there exists a repertoire of quadratic time frequency distributions that can analyze the time varying frequency content for you, ((Refer Time: 22:54)) depends on which signal you are looking at and what we have been studying is the math behind these distributions, and what freedom I have in choosing this distribution and what properties get guaranteed to have and they are not guaranteed to have and so on. So, what I suggest as an exercises try out the other distributions routines for which are available in the time frequency tool box, and see for yourself how the differ for this particular signal. You can also generate distributions for other signals as well. Good luck and I will see you in the next lecture.